

SEQUENCE LISTING

<110> The University of York
Ian Graham

<120> Desaturase Enzymes

<130> 2902076.4

<140> 10/597,998
<141> 2007-06-18

<150> 0403452.6
<151> 2004-02-17

<150> 0407807.7
<151> 2004-04-06

<160> 38

<170> PatentIn version 3.3

<210> 1
<211> 477
<212> PRT
<213> Thalassiosira pseudonana

<400> 1

Met Asp Phe Leu Ser Gly Asp Pro Phe Arg Thr Leu Val Leu Ala Ala
1 5 10 15

Leu Val Val Ile Gly Phe Ala Ala Ala Trp Gln Cys Phe Tyr Pro Pro
20 25 30

Ser Ile Val Gly Lys Pro Arg Thr Leu Ser Asn Gly Lys Leu Asn Thr
35 40 45

Arg Ile His Gly Lys Leu Tyr Asp Leu Ser Ser Phe Gln His Pro Gly
50 55 60

Gly Pro Val Ala Leu Ser Leu Val Gln Gly Arg Asp Gly Thr Ala Leu
65 70 75 80

Phe Glu Ser His His Pro Phe Ile Pro Arg Lys Asn Leu Leu Gln Ile
85 90 95

Leu Ser Lys Tyr Glu Val Pro Ser Thr Glu Asp Ser Val Ser Phe Ile
100 105 110

Ala Thr Leu Asp Glu Leu Asn Gly Glu Ser Pro Tyr Asp Trp Lys Asp
115 120 125

Ile Glu Asn Asp Asp Phe Val Ser Asp Leu Arg Ala Leu Val Ile Glu
130 135 140

His Phe Ser Pro Leu Ala Lys Glu Arg Gly Val Ser Leu Val Glu Ser
145 150 155 160

Ser Lys Ala Thr Pro Gln Arg Trp Met Val Val Leu Leu Leu Ala
Page 1

165

170

175

Ser Phe Phe Leu Ser Ile Pro Leu Tyr Leu Ser Gly Ser Trp Thr Phe
180 185 190

Val Val Val Thr Pro Ile Leu Ala Trp Leu Ala Val Val Asn Tyr Trp
195 200 205

His Asp Ala Thr His Phe Ala Leu Ser Ser Asn Trp Ile Leu Asn Ala
210 215 220

Ala Leu Pro Tyr Leu Leu Pro Leu Leu Ser Ser Pro Ser Met Trp Tyr
225 230 235 240

His His His Val Ile Gly His His Ala Tyr Thr Asn Ile Ser Lys Arg
245 250 255

Asp Pro Asp Leu Ala His Ala Pro Gln Leu Met Arg Glu His Lys Ser
260 265 270

Ile Lys Trp Arg Pro Ser His Leu Asn Gln Thr Gln Leu Pro Arg Ile
275 280 285

Leu Phe Ile Trp Ser Ile Ala Val Gly Ile Gly Leu Asn Leu Leu Asn
290 295 300

Asp Val Arg Ala Leu Thr Lys Leu Ser Tyr Asn Asn Val Val Arg Val
305 310 315 320

Glu Lys Met Ser Ser Arg Thr Leu Leu His Phe Leu Gly Arg Met
325 330 335

Leu His Ile Phe Val Thr Thr Leu Trp Pro Phe Leu Ala Phe Pro Val
340 345 350

Trp Lys Ala Ile Val Trp Ala Thr Val Pro Asn Ala Ile Leu Ser Leu
355 360 365

Cys Phe Met Leu Asn Thr Gln Ile Asn His Leu Ile Asn Thr Cys Ala
370 375 380

His Ala Ser Asp Asn Asn Phe Tyr Lys His Gln Val Val Thr Ala Gln
385 390 395 400

Asn Phe Gly Arg Ser Ser Ala Phe Cys Phe Ile Phe Ser Gly Gly Leu
405 410 415

Asn Tyr Gln Ile Glu His His Leu Leu Pro Thr Val Asn His Cys His
420 425 430

Leu Pro Ala Leu Ala Pro Gly Val Glu Arg Leu Cys Lys Lys His Gly
435 440 445

Val Thr Tyr Asn Ser Val Glu Gly Tyr Arg Glu Ala Ile Ile Ala His
450 455 460

Phe Ala His Thr Lys Asp Met Ser Thr Lys Pro Thr Asp
465 470 475

<210> 2
<211> 2914
<212> DNA
<213> Thalassiosira pseudonana

<220>
<221> misc_feature
<222> (274)..(274)
<223> n is a, c, g, or t

<400> 2
cgaggggcgc actggagagg ttccggagg ttgtatgtaa ttggagggtt ggtcaaatac 60
agattctgcc ctaacatttt ccggaaattt gcttcagittt gattcaagcg aggaggcgct 120
ccgcaggagg gcccgtcacc ttttgcata ttggggactt caatggttc tacatttttt 180
cccttctgga accccaaacgc tgtccctcaa ttctcctcc catactcacg gatggatccc 240
cgaaaatgcc acccacaatc acccttgtca atcncaaacc tcgtcatcct tcacatttt 300
ttagcaccat tggccggtgtt acccttcccc gcgactgcca gtctatgggt cagtatatct 360
cccacattt gagaggtatt gctaaaacgt gtcaatcata catatgataa ctggagagt 420
cacacgaaga gatcaatgct tgagcttagga gggtggtat tggctgttag gggcagctt 480
cacttaagat attacggcac ggcaagtcta ctgcacaata caaccgatgc tgcaaggat 540
tgcaatagct caagtgtat caacaacaaa acgtgcgaga atgacgacag tgcttacgaa 600
gatgatgcca tgagagctgt ttgggcattt ctatggcgat tgcagctggg aacgttggc 660
ggttgtgcgt tagtgttagg agtgcatcat ttcaatggag ataacctgac caaacaatct 720
gcgataccaa caaaatctc aaaagcaaag ccaatatctg atcaaaaagc agctgtgaca 780
tccggcagta cctgcgttgtt gagagagaag gcacgaaaag acggcttagt actcctcgat 840
ggcaactggt acaacgttga aaagttcgtc catcatcatc ctggagggtt agaagtgtt 900
gagcagtatc tcggggcaga tatctcgat ttgttttagag tgcgtatcatc aatccaact 960
caaatcatga aatatcgcaa gccggtacga gctgccaccc cagaagaact tgaggctctc 1020
acaagccgcc gtcaagaggt ttgtcttgat atgatggacg actttgttac caattccatt 1080
gatatcgctt ctccagaaat gctcccaag ccaacgcagt ttgacctgaa gtcatttgag 1140
aaggacttca ttgacttata tgaagagttt ttgtgtcagg gatacttcaa gcccctcaaca 1200
acatggctac tctggAACAC agcggtactg attagtatca tcgcgttatac tgcataatca 1260
atgaaaagtgc taccaccaac ttgcgttgc ctacctggag cattgcgttgg tctctttgg 1320
caccaaaagtgc gattcctcat gcacgatgcc gagcaccata atttggctgg aaacgaacgg 1380
ctgaatgaca ttgggggtt gatctatggc actgtcttct tgggtgtcaa tggcgttgg 1440
tggagagagg agcatagaga acatcatgct ttccctcaaca cttacgtatgc tggaaagtgg 1500

tcaaagatc cccaggtgt tcagcgtcac ttagacgac tc当地gatc 1560
tc当地gctca cacattcgat tt当地tattcatt cactcacaga tgagagagga cgtctggata 1620
cagaacaaga agttgattcc gttctcggt gacgagatca tt当地tctt aacaacttt 1680
cagcacattc tt当地cttcc gatcatctt atcgtggcc gc当地tggat tgc当地tagat 1740
tctacactga ctgagaggaa gttccgtcc tggagtaatgt gt当地attggg attcattgag 1800
aaggaaactgc tgat当地tact ttc当地actaa ctaactgcat cggccacttca tc当地gacgat 1860
agcaataactt gtaatgttt gtc当地atcct actacactac gcaatcttctt ctc当地agcag 1920
tc当地cttatac cccgtgtaca tcatcggctc tt当地tggcaa gctattcttctt ctttgc当地t 1980
gc当地tggaaat cactacgtca agc当地tggaa tagactcaac gatgccc当地ag aggaaactt 2040
ctgc当地ttgg cagatactaa gcaactcaaga ct当地tgc当地t ccacgtgg tctc当地ggct 2100
gtacggaggt ct当地acttcc actattccca tcatc当地ttcc ccaacgttctt ct当地agagatg 2160
ctt当地cacatt acatcaccac gc当地tgggt gaggctc当地t gtttagtgg tctacattca 2220
tatcaatgat actcatagct cc当地ttctt cgacagagac tatgtgagaa gc当地cgggctt 2280
ccg当地ttattt agattgc当地t cattgattgc gttgtggaa tggtaaccaa ct当地taacgaa 2340
gtgagggaaag acttc当地tac gaaaggccac gggagggttgg ct当地tcatgta cacgtgatct 2400
taagtgc当地a gacgatata gaggatgat ttactgttt gtc当地ccaggta gttc当地tctaa 2460
tatgatgttag caaccgc当地c ttgtggaaatt agtttagtgt actatgtaac tgaaaaagtt 2520
acgtcgatct actctctgca catctacatc gt当地tgaaggcc attccgttca agaagtatcc 2580
taatccctcg aaccaaacag tctc当地tctta taccctatcat taatcagccg cctctaccgg 2640
atgttgctgt tggc当地ggct gctgctgaaac cccctcgccg cccgataatg gc当地aaggcc 2700
gtc当地ggacact tgataatctt ct当地tacagag tt当地ttagctt ggggttgg accaataacct 2760
cctt当地tatag gtactaatgg acccggttcc attattgctt ggccgc当地t ccaccgttgg 2820
gaccgatagg tggccaaagg cccacacaga agagcaccat aaaggc当地cag ccttgaggaa 2880
actcaagaaa ccccgatggt ccacgttata aaac 2914

<210> 3
<211> 1551
<212> DNA
<213> Thalassiosira pseudonana

<400> 3
atggcttagag ctg当地ttggc attgctatgg gc当地tgc当地c tggaaacgtt ggtc当地gggtt 60
gc当地ttagtgt taggagtgca tcatc当地tctt ggagataacc tgaccaaaca atc当地tgc当地ta 120
ccaacaaaat ct当地tcaaaagc aaaggccaaa tctgatcaaa aagcagctgt gacatccggc 180
agtacctgca ct当地tgagaga gaaggcacga aaagacggc tagtactcct cgatggcaac 240
tggtaacaacg tt当地aaaagtt cgtccatcat catcctggag gt当地tagaagt gttggaggcag 300
tacctcgccgg cagataatctc gttgtgtt agagtgatgc atagaaatcc aactcaaatc 360
atgaaatatc gcaaggccggt acgagctgcc accccagaag aacttgaggc tctcacaagc 420

cgccgtcaag aggttgtct tgatatgatg gacgacttg ttaccaattc cattgatatc 480
gcttctccag aaatgctcc caagccaacg cagttgacc tgaagtcatt tgagaaggac 540
ttcattgact tataatgaaaga gtttgtgct cagggatact tcaagccctc aacaacatgg 600
ctactctgga acacagcggt actgatttagt atcatcgctg tatctgtcat ctcaatgaaa 660
gtgctaccac caacttcggt tgcctaccc ggagcattgc ttggctctt ttggcaccaa 720
agtggattcc tcatgcacga tgccgagcac cataattgg ctggaaacga acggctgaat 780
gacattttgg gttggatcta tggcactgtc ttctgggtg tcaatggcgc ttggtgaga 840
gaggagcata gagaacatca tgcttcctc aacacttacg atgatgaaag tggttcaaa 900
gatccccaga tgagagagga cgtctggata cagaacaaga agttgattcc gttctcggt 960
gacgagatca ttcatctt aacaaactt cagcacattc tggccttcc gatcatctt 1020
atcgtggcc gcgttgtat tgcgttagat tctacactga ctgagaggaa gttccgtcct 1080
tggacaatac ttggtatgt ttgtcatatc ctactacact acgcaatctt atctcagacg 1140
agtgcgtcta tccccgtgt a catcatcggt tctcttggc aagctattct ctcttgcaa 1200
ttgctggga atcactacgt caagccttgg aatagactca acgatgccac agagggaaac 1260
ttctgcgtt ggcagatact aagcactcaa gacttgcattt gtccacgtt gtctcggtgg 1320
ctgtacggag gtctcaactt tcactattcc catcatctgt tcccaacgtt gtctagagag 1380
tacttcaca ttacatcacc acgcattcgg agactatgtt agaagcacgg gttccgtt 1440
atttagatgtt cgttatgtt ttgcgttgtt ggaatggta acaacttaa cgaagtgagg 1500
aaagacttcg ctacgaaagg ccacggaggt gtggcttca tgtacacgtg a 1551

<210> 4
<211> 516
<212> PRT
<213> Thalassiosira pseudonana

<400> 4

Met Ala Arg Ala Val Trp Ala Leu Leu Trp Ala Leu Gln Leu Gly Thr
1 5 10 15

Leu Val Gly Cys Ala Leu Val Leu Gly Val His His Phe Ser Gly Asp
20 25 30

Asn Leu Thr Lys Gln Ser Ala Ile Pro Thr Lys Ser Ser Lys Ala Lys
35 40 45

Pro Ile Ser Asp Gln Lys Ala Ala Val Thr Ser Gly Ser Thr Cys Ala
50 55 60

Val Arg Glu Lys Ala Arg Lys Asp Gly Leu Val Leu Leu Asp Gly Asn
65 70 75 80

Trp Tyr Asn Val Glu Lys Phe Val His His Pro Gly Gly Val Glu
85 90 95

Val Leu Glu Gln Tyr Leu Gly Ala Asp Ile Ser Phe Val Phe Arg Val
100 105 110

Met His Arg Asn Pro Thr Gln Ile Met Lys Tyr Arg Lys Pro Val Arg
115 120 125

Ala Ala Thr Pro Glu Glu Leu Glu Ala Leu Thr Ser Arg Arg Gln Glu
130 135 140

Val Cys Leu Asp Met Met Asp Asp Phe Val Thr Asn Ser Ile Asp Ile
145 150 155 160

Ala Ser Pro Glu Met Leu Pro Lys Pro Thr Gln Phe Asp Leu Lys Ser
165 170 175

Phe Glu Lys Asp Phe Ile Asp Leu Tyr Glu Glu Phe Val Ala Gln Gly
180 185 190

Tyr Phe Lys Pro Ser Thr Thr Trp Leu Leu Trp Asn Thr Ala Val Leu
195 200 205

Ile Ser Ile Ile Ala Leu Ser Val Ile Ser Met Lys Val Leu Pro Pro
210 215 220

Thr Ser Phe Val Leu Pro Gly Ala Leu Leu Gly Leu Phe Trp His Gln
225 230 235 240

Ser Gly Phe Leu Met His Asp Ala Glu His His Asn Leu Ala Gly Asn
245 250 255

Glu Arg Leu Asn Asp Ile Leu Gly Trp Ile Tyr Gly Thr Val Phe Leu
260 265 270

Gly Val Asn Gly Ala Trp Trp Arg Glu Glu His Arg Glu His His Ala
275 280 285

Phe Leu Asn Thr Tyr Asp Asp Glu Ser Gly Phe Lys Asp Pro Gln Met
290 295 300

Arg Glu Asp Val Trp Ile Gln Asn Lys Leu Ile Pro Phe Phe Gly
305 310 315 320

Asp Glu Ile Ile His Phe Leu Thr Asn Phe Gln His Ile Leu Phe Leu
325 330 335

Pro Ile Ile Phe Ile Val Gly Arg Val Gly Ile Val Val Asp Ser Thr
340 345 350

Leu Thr Glu Arg Lys Phe Arg Pro Trp Thr Ile Leu Gly Asn Val Cys
355 360 365

His Ile Leu Leu His Tyr Ala Ile Leu Ser Gln Thr Ser Arg Pro Ile
370 375 380

Pro Val Tyr Ile Ile Gly Ser Leu Trp Gln Ala Ile Leu Ser Leu Gln
385 390 395 400

Leu Leu Gly Asn His Tyr Val Lys Pro Trp Asn Arg Leu Asn Asp Ala
405 410 415

Thr Glu Gly Asn Phe Cys Val Trp Gln Ile Leu Ser Thr Gln Asp Phe
420 425 430

Ala Cys Pro Arg Trp Ser Arg Trp Leu Tyr Gly Gly Leu Asn Phe His
435 440 445

Tyr Ser His His Leu Phe Pro Thr Leu Ser Arg Glu Tyr Phe His Ile
450 455 460

Thr Ser Pro Arg Ile Arg Arg Leu Cys Glu Lys His Gly Leu Pro Phe
465 470 475 480

Ile Glu Ile Ala Phe Ile Asp Cys Val Val Gly Met Val Asn Asn Phe
485 490 495

Asn Glu Val Arg Lys Asp Phe Ala Thr Lys Gly His Gly Ser Val Ala
500 505 510

Phe Met Tyr Thr
515

<210> 5
<211> 1900
<212> DNA
<213> Thalassiosira pseudonana

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<400> 5
nanccatatg cggaaatacg gccagggtat acccacagcg cctccgtgc agcaaactcc 60

tatccaatac ctcccatga accccccctt cgccaccctt atatgcgaga ctcgttcgtc 120

tggacctgca gatgatgact ggtgaggcca aattagttgg gaatgcgtgc agatggaggc 180

cttattcttt tgcaatcagg ggcgtcgta agaggagatc catgttgttgcgttgc 240

cttgcttggg gcgtgcatga tgtgtcggtcg cgtgtacgt gttgataggt agaaagagat 300

cgagggcggtt attcaactat tcaggatact gaaagagttt atatagcagc agtaataat 360

ccttagttgtt tgtgtttgtt ttgtgggttcaagtttcaatgcgcaaaataacgtt 420

ggttagtgtat gggtaaacag gtgttcggga caaaggctttt cttataaaatc tatttaacgtt 480

gttcgtaaa acgacgaaaa gaagccactc tgaccattc cagcgagac aagaccagca 540
 ggcacagaac agcacgacac accgacccga gccaaaaag ccaacaacaa cgacaccgac 600
 ccgagccgat acagccgaca ggcaaaggct ctctgctaca atctacaaaa cggcaacatc 660
 aaatcatgcc accctccatc aaagacacac tcgacgagcc cttcgctcg cccgcatcca 720
 ccaagtgcgc caccaccaaa cccctccccc cccgcccggaa accccctaaaa cgatactccc 780
 cctccaaat ctcccaacac aacactccca ccgatgcatg gtcatttac aaatcccaag 840
 tcctgacat ttccaaatgg atatgcacc atccagggtgg agagcagacg ctgttgaggt 900
 ttgccccgtt ggtatgtacc gatgaattga gggcattca tgatgttgg gttttggagg 960
 agaagtgcgc tcattttgtt attggggagg tggattggac tactaccggc ggggcagaga 1020
 atactgtcac gaaggatgga cagggttcgg agcttatcaa ggattcaga gagttgggtg 1080
 aacacttcga cagggtgggg tactttcagc tcagttccatg gtattacgtc cgtaaagggtgg 1140
 ctaccgttccatccatggatgtgcac tcggactcctt cttcaatacc gattccatcc 1200
 cagcacacat gctcgccggcg gtactcctcg gtatattctg gcaacaattt gcattcgctcg 1260
 gacatgactg tggtcacatg tcggcgccggaa ctcatgccccg tgatcatatc gatgtaccta 1320
 agctggggagc actggtgacc ttcttcaatg ggatttcgggt agcgtgggtgg aaggctacgc 1380
 acaatgttca tcatgctgtg ccaaatagtg ttgattgtga cccggacatt gtcatttgc 1440
 cggtgttgc gttgcatgag cacatgtta cgtcggttttta taacaagtttcatgggaggg 1500
 tggatgggtt tggatggctg ggcgtatgttgc ttcttgcatttcaacac ttttgggtact 1560
 atcccaataat ggcgggtggcg aggttcaatc tgtacattca atcagcatttgc ttttggcgt 1620
 cgaagaacga tggcatgca ggaagaaggg gatcccttag attggatttgc tggcggtca 1680
 atcgtgttct tctgtgggtt agcgggtctg gtgtcatgca tcccgagctg ggcggagcgt 1740
 atcgcatttcg tcttcgtcag acatgctgttac cctgggttac tgcatgtgca atcaccgtc 1800
 gccttcctt ggacaatctt gatcccacaa gaggaccggg ttgggggtcttccgaag 1860
 cccgggttctg ggctttgcc acattggcgt cccgggtcca 1900

<210> 6
 <211> 411
 <212> PRT
 <213> Thalassiosira pseudonana

<400> 6

Met Pro Pro Ser Ile Lys Asp Thr Leu Asp Glu Pro Phe Val Ser Pro
 1 5 10 15

Ala Ser Thr Lys Ser Pro Thr Thr Lys Pro Leu Leu Pro Arg Arg Lys
 20 25 30

Pro Leu Lys Arg Tyr Ser Pro Ser Gln Ile Ser Gln His Asn Thr Pro
 35 40 45

Thr Asp Ala Trp Leu Ile Tyr Lys Ser Gln Val Leu Asp Ile Ser Lys
 Page 8

50 55 60

Trp Ile Ser His His Pro Gly Gly Glu Gln Thr Leu Leu Arg Phe Ala
65 70 75 80

Gly Met Asp Ala Thr Asp Glu Leu Arg Ala Phe His Asp Asp Trp Val
85 90 95

Leu Glu Glu Lys Leu Pro His Phe Val Ile Gly Glu Val Asp Trp Thr
100 105 110

Thr Thr Gly Gly Ala Glu Asn Thr Val Thr Lys Asp Gly Gln Val Ser
115 120 125

Glu Leu Ile Lys Asp Phe Arg Glu Leu Gly Glu His Phe Asp Arg Leu
130 135 140

Gly Tyr Phe His Val Ser Pro Trp Tyr Tyr Val Arg Lys Val Ala Thr
145 150 155 160

Val Phe Ala Ile Phe Gly Cys Ala Leu Gly Leu Leu Phe Asn Thr Asp
165 170 175

Ser Ile Pro Ala His Met Leu Ala Ala Val Leu Leu Gly Ile Phe Trp
180 185 190

Gln Gln Phe Ala Phe Val Gly His Asp Cys Gly His Met Ser Ala Arg
195 200 205

Thr His Ala Arg Asp His Ile Asp Val Pro Lys Leu Gly Ala Leu Val
210 215 220

Thr Phe Phe Asn Gly Ile Ser Val Ala Trp Trp Lys Ala Thr His Asn
225 230 235 240

Val His His Ala Val Pro Asn Ser Val Asp Cys Asp Pro Asp Ile Ala
245 250 255

His Leu Pro Val Phe Ala Leu His Glu His Met Phe Thr Ser Leu Phe
260 265 270

Asn Lys Tyr His Gly Arg Val Met Glu Phe Asp Trp Leu Ala Arg Asn
275 280 285

Val Phe Val Pro Phe Gln His Phe Trp Tyr Tyr Pro Ile Met Ala Val
290 295 300

Ala Arg Phe Asn Leu Tyr Ile Gln Ser Ala Leu Phe Leu Ala Ser Lys
305 310 315 320

Asn Asp Gly His Ala Gly Arg Arg Gly Ser Ser Arg Leu Asp Leu Leu
325 330 335

Ala Phe Asn Arg Val Leu Leu Leu Val Ser Gly Ala Gly Val Met His
340 345 350

Pro Glu Leu Gly Gly Ala Tyr Arg Ile Arg Leu Arg Gln Thr Cys Cys
355 360 365

Thr Trp Val Thr Ala Cys Ala Ile Thr Cys Arg Leu Leu Leu Asp Asn
370 375 380

Leu Asp Pro Thr Arg Gly Pro Gly Trp Gly Ala Leu Ser Glu Ala Arg
385 390 395 400

Phe Trp Ala Phe Ala Thr Leu Ala Ser Arg Val
405 410

<210> 7
<211> 1913
<212> DNA
<213> Thalassiosira pseudonana

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1800)..(1800)
<223> n is a, c, g, or t

<400> 7
cancataaccg ggaagagggc cttatttgcc accacagtga taaccttcgg ctgtgaccac 60
gggagcagcc gtggcgagcc cgcgctgac cagccctgtc ttttggagc atccctcacc 120
acacatcgca tctcgltgca cggggatcag tgcacagtct tcgtctcatt gtttagatgta 180
cacgcgaaga agcacatcca gcccgaactct tcataacatc tcaggaccct gcaaacacgc 240
atcacatcat gatgttccac cgagtcgtca tcggcatcgc ctcacaatg ggctgtgtct 300
ccagttctc ctgc(cccggt cattcaaatat tggcacgtcc tatgcaatca tccaccactt 360
ctcglttctc gacaatgatt gaaaagtca gatgttctgca cagtgtaac aacaaaaaca 420
aggagatgac atcatcttct gaaatgccta ctgcgtggga atgcaatgag gaagctgagt 480
gcgtggaagt tcctgctgt gatgacgagg aatgccgtac tacttggat gtgaggattc 540
atggcaaatg gtacgatctt tcaggtgagt gcaagttgtg gtatgcattt ttataaggttc 600
tattctgtat cggcacacac gatattgtgt tttgtatcaat gttctaacag ccatttttttc 660
ctcctacttc ctcaggatgg cgcaaagctc accctgcagg accccactgg atcgactgg 720
acgacggtcg tgacgccacc gaagtcatgg acgcattca cacccaaaaa ggacgtgaaa 780
tgtacaagcg tctcccgcg tctgcccccg aaacggctgc cgttctgaa gcatctgcag 840
caccttactc gcagacggag cttaaacttta ggaagttgag ggatcaattg gaaagtgagg 900
ggtgtggaa gaggacttt gtccatgagg gaaagttgct ggcgatttg gcatcggttgg 960

ttacaggagc agcattgact gcggagagtg ctcctccctt ttcaacttc ttgttggat 1020
 tgtctatgac gaatgctgga tggggggc atgattatat tcatgggtt gataagtca 1080
 gtcaagttat gaggccttt gctgccgtgg ctgctggtt gggaccaact tggtgagtg 1140
 ataagcacaa caagcatcac gcttgagtg agtctgactc ttgttgcac tgcaagtgt 1200
 gttaaagat tgaatcaata ccatcgact catatcctca acattcttc aatcgcaaca 1260
 gccaacgaaa tggagttga tgaagacatt gcgaccgatc catttcctt tccttatgtc 1320
 ccggatccaa agtacgattc tccacttcgt aagatccaac actacatctt ctacagtccc 1380
 ttctccctcc tccttgccct ctggcgctg gacaccctta aggtcgccgt agactcagtt 1440
 gaatcgaaac gtcccgatgc aaagaatgaa ttgtggtac tcttggcaca ttacttcgtc 1500
 ttgtgacct tcctccaggc tcaggtgtgg gtgcctgctg tcctcccttc tggcctcatg 1560
 tctgcactca ttgttactcc gacacatcag tcggaagagt atttttagga gtatcagcct 1620
 gattgggtga cggttcagtt tgagagcacg agaaatgctg tcacgactaa tccattctct 1680
 gagtggctt ggggaggaat gcaataccag ttggagcatc acttggccc ttccatgccc 1740
 aggttaagcag cttaatgtt gtatcttgc ccatgttgc ctctcggtc tcggctaacn 1800
 ctgttggaaag cgtatgagcc tagcacataa tggtggtat gcgaccatga actcgattta 1860
 aggtcaaat accttactat catctcagtc cggtgcccga tgacgtgtt ccc 1913

<210> 8
 <211> 389
 <212> PRT
 <213> Thalassiosira pseudonana

<400> 8

Gln Pro Phe Val Pro Pro Thr Ser Ser Gly Trp Arg Lys Ala His Pro
 1 5 10 15

Ala Gly Pro His Trp Ile Asp Trp Tyr Asp Gly Arg Asp Ala Thr Glu
 20 25 30

Val Met Asp Ala Phe His Thr Gln Lys Gly Arg Glu Met Tyr Lys Arg
 35 40 45

Leu Pro Ala Ser Ala Pro Glu Thr Ala Ala Val Leu Glu Ala Ser Ala
 50 55 60

Ala Pro Tyr Ser Gln Thr Glu Leu Asn Phe Arg Lys Leu Arg Asp Gln
 65 70 75 80

Leu Glu Ser Glu Gly Trp Trp Glu Arg Asp Phe Val His Glu Gly Lys
 85 90 95

Leu Leu Ala Ile Trp Ala Ser Leu Val Thr Gly Ala Ala Leu Thr Ala
 100 105 110

Glu Ser Ala Pro Pro Leu Ser Thr Phe Leu Leu Gly Leu Ser Met Thr
 115 120 125

Asn Ala Gly Trp Leu Gly His Asp Tyr Ile His Gly Val Asp Lys Phe
130 135 140

Ser Gln Val Met Arg Pro Phe Ala Ala Val Ala Ala Gly Leu Gly Pro
145 150 155 160

Thr Trp Trp Ser Asp Lys His Asn Lys His His Ala Leu Ser Glu Ser
165 170 175

Asp Ser Cys Cys Tyr Cys Lys Cys Gly Leu Lys Ile Glu Ser Ile Pro
180 185 190

Ser Tyr Ser Tyr Pro Gln His Ser Phe Asn Arg Asn Ser Gln Arg Asn
195 200 205

Gly Ser Arg Leu Asn Gln Tyr His Arg Thr His Ile Leu Asn Ile Leu
210 215 220

Ser Ile Ala Thr Ala Asn Glu Met Gly Val Asp Glu Asp Ile Ala Thr
225 230 235 240

Asp Pro Phe Leu Phe Pro Tyr Val Pro Asp Pro Lys Tyr Asp Ser Pro
245 250 255

Leu Arg Lys Ile Gln His Tyr Ile Phe Tyr Ser Pro Phe Ser Phe Leu
260 265 270

Phe Ala Leu Trp Arg Val Asp Thr Leu Lys Val Ala Val Asp Ser Val
275 280 285

Glu Ser Lys Arg Pro Asp Ala Lys Asn Glu Leu Trp Tyr Leu Leu Ala
290 295 300

His Tyr Phe Val Leu Leu Thr Phe Phe Pro Ala Gln Val Trp Val Pro
305 310 315 320

Ala Val Phe Leu Ser Gly Leu Met Ser Ala Leu Ile Val Thr Pro Thr
325 330 335

His Gln Ser Glu Glu Tyr Phe Glu Glu Tyr Gln Pro Asp Trp Val Thr
340 345 350

Ala Gln Phe Glu Ser Thr Arg Asn Ala Val Thr Thr Asn Pro Phe Ser
355 360 365

Glu Trp Leu Trp Gly Gly Met Gln Tyr Gln Leu Glu His His Leu Phe
370 375 380

Pro Ser Met Pro Arg
385

<210> 9
<211> 2689
<212> DNA
<213> Thalassiosira pseudonana

<220>
<221> misc_feature
<222> (14)..(17)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2687)..(2687)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2689)..(2689)
<223> n is a, c, g, or t

<400> 9
aaaaaaaaaa aaannnnnnn aagcgagatc aatcgagctg gtaccatgag tttcaaaagt 60
caacttcaac attcaagttg tacaaaagag agggcctcag acgtggtgag caaaagcact 120
tcacaggggaa atatgggg aaaaacagaa atatggca aatttatctt agtccctgat 180
tatatcttca attactaaag gaaaaacaat gcagctcaa agctacgtt gtgtacttct 240
ttgaaaccac ctcacccccc cggcttcgcg tccgggtcgg cccgcctgca tccttcttc 300
ctctcacaat ttatcatcca acgagctgat aacgtgtcat ttcacagggt caacacaata 360
aaacatacta atcaaccatg gaaaaaggag gagacgcagc cgcagccacc aagcgtagtg 420
gagcattgaa attggcggag aagccgcaga agtacacctg gcaggaggtg aagaagcacg 480
tgagtctccg cttgtgtgc tgccgttgg a tgccttgcgttgcgg attatgcaac 540
gagagttcgt attgcaactc aatttcaattt gtccatctgc aatcaactca tctgacccaa 600
caacttgc caccgtccac ccattcagat caccggac gatgcctgg tagtccacca 660
aaacaaaagtc tacgacgtct ccaactggta cgaccacccc ggtggagccg tgggttac 720
ccacgcccga gacgacatga cggacatctt cggcccttc cacgcccag gctctcaggc 780
catgatgaag aagtttaca ttggagattt gattccggag agtgtggagc ataaggatca 840
aagacagttg gatttcgaga agggataatcg tgattacgg gccaagcttg tcatgatggg 900
gatgtcaag tcgagtaaga tgtattatgc atacaagtgc tcgttcaata tgtgcatttg 960
gttgggtggcg gtggccatgg tgtaatctc ggacagttt gcaatgcaca ttggatccgc 1020
tctctgttgc ggattgttgc ggcacgcgtg tggatggctt ggcacgcact ttcttcacca 1080
ccaagtctt aagcaacgaa agtacggaga tctcggttgc atcttttggg gagatctcat 1140
gcaggggttc tcgatgcagt ggtggaaagaa caagcacaat ggcacccatg ctgtcccaa 1200
cttgacacaac tttcccttgg acagtcagga tggatggcc gatattgata ccatgccact 1260
ccttgcttgg agtctcaagc aggctcagag tttcagagag atcaataagg gaaaggacag 1320
tacttcgta aagtacgcta tcaaattcca ggcattcaca tactccccca tcctcctt 1380
ggctcgcatc tcttgggttga atgaatcctt caaaaactgca ttggactcg gagctgcctc 1440

ggagaatgcc aagtggagt tggagaagcg tggacttcag taccactt tggagaagct 1500
 tggaaatcacc ctccattaca ctggatgtt cgtccctct tccggatttgaagggtggtc 1560
 tctccatat tccatcatgt atttcttcac tgccacatgc tcctcggac tttcctcgc 1620
 atggcttt ggattggac acaacggtat gtcagtgtac gatgccacca cccgacctga 1680
 cttctggcaa ctccaagtca ccactacacg taacatcatt ggtggacacg gcattcccc 1740
 attcttgc gattggttct gcgggtggatt gcaataccaa gtggatcacc acctctccc 1800
 catgatgcct agaaaacaata tcgcaaatg ccacaagctt gtggagtc tctgtaagga 1860
 gtgggggttg aagtaccatg aggccgatat gtggatggt accgtggaaatgttgcaca 1920
 tctctccaag gtgtcgatg atttcctgtt ggagatgggt aaggatttcc ctgcatgt 1980
 aacaccttatt accagtccgc agctttgtcg gttgctggag atgaatgtatc cgaactcatc 2040
 gtaaaatactc attattaatg aacaatgtta ccctgcagtc gtgagggttgcctcgtt 2100
 cccacccctt ctattgtgtt ttggatgtca ttgaaacgag atagtcttatttctacatc 2160
 atctctccat tcaccctcga atagtatccc aacaaccatc acatcaaact acttgaatct 2220
 cctctgtggc aatccctccc attgtacatt tactctcaaa ggtatctta ttgtccctt 2280
 tattaaattgt tgaatattga agggaaagat tccattttcc cctctctt ccccgatgt 2340
 cctctcacct ctaaatacct ttacacaacac aacaacgaaa caacgcagat cagacaaaca 2400
 acatggcaga actatcctca ccgtgcaaac gatccaaagg cgaagagcta ttcttagtcc 2460
 atctccaacg catgtctggc tccagaccct catcctgaag agtggatgtt gatgtcgctg 2520
 atgtacttgc cgtctgtat ttctctgagg tgcacaact cagggtcacc aaagcagctt 2580
 cgctgatcgc tagtggcgag aagatccat ttccatccc gaagaaagcc tccctggaaa 2640
 aatgtcaactt ctgaaagtc gagggtgacg aataattggg ggcggangn 2689

<210> 10
 <211> 1458
 <212> PRT
 <213> Thalassiosira pseudonana

<400> 10

Ala Thr Gly Gly Cys Thr Gly Gly Ala Ala Ala Ala Gly Gly Ala Gly
 1 5 10 15

Gly Ala Gly Ala Cys Gly Cys Ala Gly Cys Cys Gly Cys Ala Gly Cys
 20 25 30

Thr Ala Cys Cys Ala Ala Gly Cys Gly Thr Ala Gly Thr Gly Gly Ala
 35 40 45

Gly Cys Ala Thr Thr Gly Ala Ala Ala Thr Thr Gly Gly Cys Gly Gly
 50 55 60

Ala Gly Ala Ala Gly Cys Cys Gly Cys Ala Gly Ala Ala Gly Thr Ala
 65 70 75 80

Cys Ala Cys Thr Thr Gly Gly Cys Ala Gly Gly Ala Gly Thr Gly
85 90 95

Ala Ala Gly Ala Ala Gly Cys Ala Cys Ala Thr Cys Ala Cys Cys Cys
100 105 110

Cys Cys Gly Ala Cys Gly Ala Thr Gly Cys Cys Thr Gly Gly Thr
115 120 125

Ala Gly Thr Cys Cys Ala Cys Cys Ala Ala Ala Ala Cys Ala Ala Ala
130 135 140

Gly Thr Cys Thr Ala Cys Gly Ala Cys Gly Thr Cys Thr Cys Cys Ala
145 150 155 160

Ala Cys Thr Gly Gly Thr Ala Cys Gly Ala Cys Cys Ala Cys Cys Cys
165 170 175

Cys Gly Gly Thr Gly Gly Ala Gly Cys Cys Gly Thr Gly Gly Thr Gly
180 185 190

Thr Thr Cys Ala Cys Cys Cys Ala Cys Gly Cys Cys Gly Gly Ala Gly
195 200 205

Ala Cys Gly Ala Cys Ala Thr Gly Ala Cys Gly Gly Ala Cys Ala Thr
210 215 220

Cys Thr Thr Cys Gly Cys Cys Gly Cys Cys Thr Thr Cys Cys Ala Cys
225 230 235 240

Gly Cys Cys Cys Ala Ala Gly Gly Cys Thr Cys Thr Cys Ala Gly Gly
245 250 255

Cys Cys Ala Thr Gly Ala Thr Gly Ala Ala Gly Ala Ala Gly Thr Thr
260 265 270

Thr Thr Ala Cys Ala Thr Thr Gly Gly Ala Gly Ala Thr Thr Thr Gly
275 280 285

Ala Thr Thr Cys Cys Gly Gly Ala Gly Ala Gly Thr Gly Thr Gly Gly
290 295 300

Ala Gly Cys Ala Thr Ala Ala Gly Gly Ala Thr Cys Ala Ala Ala Gly
305 310 315 320

Ala Cys Ala Gly Thr Thr Gly Gly Ala Thr Thr Cys Gly Ala Gly
325 330 335

Ala Ala Gly Gly Ala Thr Ala Thr Cys Gly Thr Gly Ala Thr Thr
340 345 350

Thr Ala Cys Gly Gly Gly Cys Cys Ala Ala Gly Cys Thr Thr Gly Thr
355 360 365

Cys Ala Thr Gly Ala Thr Gly Gly Gly Ala Thr Gly Thr Thr Cys
370 375 380

Ala Ala Gly Thr Cys Gly Ala Gly Thr Ala Ala Gly Ala Thr Gly Thr
385 390 395 400

Ala Thr Thr Ala Thr Gly Cys Ala Thr Ala Cys Ala Ala Gly Thr Gly
405 410 415

Cys Thr Cys Gly Thr Thr Cys Ala Ala Thr Ala Thr Gly Thr Gly Cys
420 425 430

Ala Thr Gly Thr Gly Gly Thr Thr Gly Gly Thr Gly Cys Gly Gly
435 440 445

Thr Gly Gly Cys Cys Ala Thr Gly Gly Thr Gly Thr Ala Cys Thr Ala
450 455 460

Cys Thr Cys Gly Gly Ala Cys Ala Gly Thr Thr Thr Gly Gly Cys Ala
465 470 475 480

Ala Thr Gly Cys Ala Cys Ala Thr Thr Gly Gly Ala Thr Cys Gly Gly
485 490 495

Cys Thr Cys Thr Cys Thr Thr Gly Thr Thr Gly Gly Ala Thr Thr
500 505 510

Gly Thr Thr Cys Thr Gly Gly Cys Ala Gly Cys Ala Gly Thr Gly Thr
515 520 525

Gly Gly Ala Thr Gly Gly Cys Thr Thr Gly Cys Gly Cys Ala Cys Gly
530 535 540

Ala Cys Thr Thr Thr Cys Thr Thr Cys Ala Cys Cys Ala Cys Cys Ala
545 550 555 560

Ala Gly Thr Cys Thr Thr Ala Ala Gly Cys Ala Ala Cys Gly Ala
565 570 575

Ala Ala Gly Thr Ala Cys Gly Gly Ala Gly Ala Thr Cys Thr Cys Gly
580 585 590

Thr Thr Gly Gly Cys Ala Thr Cys Thr Thr Thr Gly Gly Gly
595 600 605

Ala Gly Ala Thr Cys Thr Cys Ala Thr Gly Cys Ala Gly Gly Gly Gly
610 615 620

Thr Thr Cys Thr Cys Gly Ala Thr Gly Cys Ala Gly Thr Gly Gly Thr
625 630 635 640

Gly Gly Ala Ala Gly Ala Ala Cys Ala Ala Gly Cys Ala Cys Ala Ala
645 650 655

Thr Gly Gly Cys Cys Ala Cys Cys Ala Thr Gly Cys Thr Gly Thr Thr
660 665 670

Cys Cys Cys Ala Ala Cys Thr Thr Gly Cys Ala Cys Ala Ala Cys Thr
675 680 685

Cys Thr Thr Cys Cys Thr Thr Gly Gly Ala Cys Ala Gly Thr Cys Ala
690 695 700

Gly Gly Ala Thr Gly Gly Thr Gly Ala Thr Cys Cys Cys Gly Ala Thr
705 710 715 720

Ala Thr Thr Gly Ala Thr Ala Cys Cys Ala Thr Gly Cys Cys Ala Cys
725 730 735

Thr Cys Cys Thr Thr Gly Cys Thr Thr Gly Gly Ala Gly Thr Cys Thr
740 745 750

Cys Ala Ala Gly Cys Ala Gly Gly Cys Thr Cys Ala Gly Ala Gly Thr
755 760 765

Thr Thr Cys Ala Gly Ala Gly Ala Thr Cys Ala Ala Thr Ala
770 775 780

Ala Gly Gly Ala Ala Ala Gly Gly Ala Cys Ala Gly Thr Ala Cys
785 790 795 800

Cys Thr Thr Cys Gly Thr Cys Ala Ala Gly Thr Ala Cys Gly Cys Thr
805 810 815

Ala Thr Cys Ala Ala Ala Thr Thr Cys Cys Ala Gly Gly Cys Ala Thr
820 825 830

Thr Cys Ala Cys Ala Thr Ala Cys Thr Thr Cys Cys Cys Ala Thr
835 840 845

Cys Cys Thr Cys Cys Thr Cys Thr Thr Gly Gly Cys Thr Cys Gly Cys
850 855 860

Ala Thr Cys Thr Cys Thr Thr Gly Gly Thr Thr Gly Ala Ala Thr Gly
865 870 875 880

Ala Ala Thr Cys Cys Thr Thr Cys Ala Ala Ala Cys Thr Gly Cys
885 890 895

Ala Thr Thr Cys Gly Gly Ala Cys Thr Cys Gly Gly Ala Gly Cys Thr
900 905 910

Gly Cys Cys Thr Cys Gly Gly Ala Gly Ala Ala Thr Gly Cys Cys Ala
915 920 925

Ala Gly Thr Thr Gly Gly Ala Gly Thr Thr Gly Gly Ala Gly Ala Ala
930 935 940

Gly Cys Gly Thr Gly Gly Ala Cys Thr Thr Cys Ala Gly Thr Ala Cys
945 950 955 960

Cys Cys Ala Cys Thr Thr Thr Gly Gly Ala Gly Ala Ala Gly Cys
965 970 975

Thr Thr Gly Gly Ala Ala Thr Cys Ala Cys Cys Cys Thr Thr Cys Ala
980 985 990

Cys Thr Ala Cys Ala Cys Thr Thr Gly Gly Ala Thr Gly Thr Thr Cys
995 1000 1005

Gly Thr Cys Cys Thr Cys Thr Cys Thr Cys Cys Gly Gly Ala
1010 1015 1020

Thr Thr Thr Gly Gly Ala Ala Gly Gly Thr Gly Gly Thr Cys Thr
1025 1030 1035

Cys Thr Thr Cys Cys Ala Thr Ala Thr Thr Cys Cys Ala Thr Cys
1040 1045 1050

Ala Thr Gly Thr Ala Thr Thr Thr Cys Thr Thr Cys Ala Cys Thr
1055 1060 1065

Gly Cys Cys Ala Cys Ala Thr Gly Cys Thr Cys Cys Thr Cys Gly
1070 1075 1080

Gly Gly Ala Cys Thr Thr Thr Thr Cys Cys Thr Cys Gly Cys Ala
1085 1090 1095

Thr Thr Gly Gly Thr Cys Thr Thr Thr Gly Gly Ala Thr Thr Gly
1100 1105 1110

Gly Gly Ala Cys Ala Cys Ala Ala Cys Gly Gly Thr Ala Thr Gly
1115 1120 1125

Thr Cys Ala Gly Thr Gly Thr Ala Cys Gly Ala Thr Gly Cys Cys
1130 1135 1140

Ala Cys Cys Ala Cys Cys Cys Gly Ala Cys Cys Thr Gly Ala Cys
1145 1150 1155

Thr Thr Cys Thr Gly Gly Cys Ala Ala Cys Thr Cys Cys Ala Ala
1160 1165 1170

Gly Thr Cys Ala Cys Cys Ala Cys Thr Ala Cys Ala Cys Gly Thr
1175 1180 1185

Ala Ala Cys Ala Thr Cys Ala Thr Thr Gly Gly Thr Gly Gly Ala
1190 1195 1200

Cys Ala Cys Gly Gly Cys Ala Thr Thr Cys Cys Cys Cys Ala Ala
1205 1210 1215

Thr Thr Cys Thr Thr Thr Gly Thr Gly Gly Ala Thr Thr Gly Gly
1220 1225 1230

Thr Thr Cys Thr Gly Cys Gly Gly Thr Gly Gly Ala Thr Thr Gly
1235 1240 1245

Cys Ala Ala Thr Ala Cys Cys Ala Ala Gly Thr Gly Gly Ala Thr
1250 1255 1260

Cys Ala Cys Cys Ala Cys Cys Thr Cys Thr Thr Cys Cys Cys Cys
1265 1270 1275

Ala Thr Gly Ala Thr Gly Cys Cys Thr Ala Gly Ala Ala Ala Cys
1280 1285 1290

Ala Ala Thr Ala Thr Cys Gly Cys Gly Ala Ala Gly Thr Gly Cys
1295 1300 1305

Cys Ala Cys Ala Ala Gly Cys Thr Thr Gly Thr Gly Gly Ala Gly
1310 1315 1320

Thr Cys Ala Thr Thr Cys Thr Gly Thr Ala Ala Gly Gly Ala Gly
1325 1330 1335

Thr Gly Gly Gly Thr Gly Thr Gly Ala Ala Gly Thr Ala Cys
1340 1345 1350

Cys Ala Thr Gly Ala Gly Gly Cys Thr Gly Ala Thr Ala Thr Gly
1355 1360 1365

Thr Gly Gly Gly Ala Thr Gly Gly Thr Ala Cys Cys Gly Thr Gly
1370 1375 1380

Gly Ala Ala Gly Thr Gly Thr Gly Cys Ala Ala Cys Ala Thr
1385 1390 1395

Cys Thr Cys Thr Cys Cys Ala Ala Gly Gly Thr Gly Thr Cys Gly
1400 1405 1410

Gly Ala Thr Gly Ala Thr Thr Thr Cys Cys Thr Thr Gly Thr Gly
1415 1420 1425

Gly Ala Gly Ala Thr Gly Gly Thr Gly Ala Ala Gly Gly Ala Thr
1430 1435 1440

Thr Thr Cys Cys Cys Thr Gly Cys Cys Ala Thr Gly Thr Ala Ala
1445 1450 1455

<210> 11
<211> 485
<212> PRT
<213> Thalassiosira pseudonana

<400> 11

Met Ala Gly Lys Gly Gly Asp Ala Ala Ala Ala Thr Lys Arg Ser Gly
1 5 10 15

Ala Leu Lys Leu Ala Glu Lys Pro Gln Lys Tyr Thr Trp Gln Glu Val
20 25 30

Lys Lys His Ile Thr Pro Asp Asp Ala Trp Val Val His Gln Asn Lys
35 40 45

Val Tyr Asp Val Ser Asn Trp Tyr Asp His Pro Gly Gly Ala Val Val
50 55 60

Phe Thr His Ala Gly Asp Asp Met Thr Asp Ile Phe Ala Ala Phe His
65 70 75 80

Ala Gln Gly Ser Gln Ala Met Met Lys Lys Phe Tyr Ile Gly Asp Leu
85 90 95

Ile Pro Glu Ser Val Glu His Lys Asp Gln Arg Gln Leu Asp Phe Glu
100 105 110

Lys Gly Tyr Arg Asp Leu Arg Ala Lys Leu Val Met Met Gly Met Phe
115 120 125

Lys Ser Ser Lys Met Tyr Tyr Ala Tyr Lys Cys Ser Phe Asn Met Cys
130 135 140

Met Trp Leu Val Ala Val Ala Met Val Tyr Tyr Ser Asp Ser Leu Ala
145 150 155 160

Met His Ile Gly Ser Ala Leu Leu Leu Gly Leu Phe Trp Gln Gln Cys
165 170 175

Gly Trp Leu Ala His Asp Phe Leu His His Gln Val Phe Lys Gln Arg
180 185 190

Lys Tyr Gly Asp Leu Val Gly Ile Phe Trp Gly Asp Leu Met Gln Gly
195 200 205

Phe Ser Met Gln Trp Trp Lys Asn Lys His Asn Gly His His Ala Val
210 215 220

Pro Asn Leu His Asn Ser Ser Leu Asp Ser Gln Asp Gly Asp Pro Asp
225 230 235 240

Ile Asp Thr Met Pro Leu Leu Ala Trp Ser Leu Lys Gln Ala Gln Ser

245

250

255

Phe Arg Glu Ile Asn Lys Gly Lys Asp Ser Thr Phe Val Lys Tyr Ala
260 265 270

Ile Lys Phe Gln Ala Phe Thr Tyr Phe Pro Ile Leu Leu Leu Ala Arg
275 280 285

Ile Ser Trp Leu Asn Glu Ser Phe Lys Thr Ala Phe Gly Leu Gly Ala
290 295 300

Ala Ser Glu Asn Ala Lys Leu Glu Leu Glu Lys Arg Gly Leu Gln Tyr
305 310 315 320

Pro Leu Leu Glu Lys Leu Gly Ile Thr Leu His Tyr Thr Trp Met Phe
325 330 335

Val Leu Ser Ser Gly Phe Gly Arg Trp Ser Leu Pro Tyr Ser Ile Met
340 345 350

Tyr Phe Phe Thr Ala Thr Cys Ser Ser Gly Leu Phe Leu Ala Leu Val
355 360 365

Phe Gly Leu Gly His Asn Gly Met Ser Val Tyr Asp Ala Thr Thr Arg
370 375 380

Pro Asp Phe Trp Gln Leu Gln Val Thr Thr Arg Asn Ile Ile Gly
385 390 395 400

Gly His Gly Ile Pro Gln Phe Phe Val Asp Trp Phe Cys Gly Gly Leu
405 410 415

Gln Tyr Gln Val Asp His His Leu Phe Pro Met Met Pro Arg Asn Asn
420 425 430

Ile Ala Lys Cys His Lys Leu Val Glu Ser Phe Cys Lys Glu Trp Gly
435 440 445

Val Lys Tyr His Glu Ala Asp Met Trp Asp Gly Thr Val Glu Val Leu
450 455 460

Gln His Leu Ser Lys Val Ser Asp Asp Phe Leu Val Glu Met Val Lys
465 470 475 480

Asp Phe Pro Ala Met
485

<210> 12
<211> 2775
<212> DNA
<213> Thalassiosira pseudonana

<220>

<221> misc_feature
<222> (2530)..(2530)
<223> n is a, c, g, or t

```
<220>
<221> misc_feature
<222> (2743)..(2743)
<223> n is a, c, g, or t
```

<400> 12
tatgtccacc ccccccgtt ttgtccacct ctgtctcgta tcttgggacc cgggtctcgta 60
gtttgcgaga cctctcaagc gggcccatag tagacgactt gatctgttg ctgatactg 120
acgtgcacccg attttcggg gctaacgcca ctttctaa ctccaccagg tacgactgac 180
ttgtccccgt agatatctt gataccctta tggcaaagcc gatcaaatcg aaatgattt 240
actgttagcaa ggataaagcag atggataggc gggggatctt catgtcgaca agaggaagag 300
agagagtatg tcgtcggcga ggggtggatag gttgagagag agggggatgac agattgtaca 360
ttatcttccc tccaagactt taccaaggca cgtcactctg attagaatct tacatacacg 420
tggagtaata gtggacaata aatgacaagt gaagcacccc agtggaccat ttctcgcca 480
cgtggtcgtc cgctgtgggt tgagtgaacc gacgacgacg aacacaaccc ctgaatctcc 540
ttcggcaaca acaatacacc aatatgtca acggcaacct cccagcatcc accgcacagc 600
tcaagtccac ctgcgaagccc cagcagcaac atgagcatcg caccatctcc aagtccgagc 660
tcgcccaca caacacgccc aaatcagcat ggtgtccgt ccactccact ccccccaccc 720
accatccca ctccaacaac aaacaacacg cacacctagt cctcgacatt accgactttg 780
cgtcccgcca tccaggggaa gacctcatcc tcctcgctc cggcaaagac gcctcggtc 840
tgttggaaac ataccatcca cgtggagttc cgacgtctt cattcaaaag ctgcagattt 900
gagtgtatgga ggaggaggcg ttccggatt cgttttacag ttggactgat tctgactttt 960
atactgtgtt gaagaggagg gtttggagc ggttggagga gaggggggtt gcgaggagg 1020
gatcgaaaga gattggatc aaggcttgc tcttgggtt tggattttgg tactgtttgt 1080
acaagatgtt tactacgtcg gatattgtatc agtacggat tgccattgcc tattctattt 1140
gaatggaaac ctttgcggca ttcatcgca cgtgtattca acacgatgga aatcacgggt 1200
cattcgctca gaacaagtta ctcaacaagt tggctgggtt gacgtggat atgattgggt 1260
cgagtgcgtt tacgtggag cttcagcaca tgctgggca tcatccatatacgaatgtgt 1320
tggatgggggt ggaggaggag aggaaggaga ggggggagga tggcttttgaagaaaagg 1380
atcaggtgag acgagatgac agagagagag agagtctatt cgtgtgaagt cgttagatgca 1440
tgtgtgcgtt tgagcgacac aactctaacg cattgcattt cacttcaac tcggcgacag 1500
gaatcagatc cagacgtatt ctcccttc cctctcatga gaatgcattcc cctccatatac 1560
acctcatgtt atcataaaata ccaacaccc tacgctccac ccctcttgc attgatgaca 1620
cttgcggaaatg tattccaaaca ggattttgaa gttgccacat ccggacgatt atatcatatt 1680
gtatggccaaatg tacgttatgg ttccggatgg aatgtcatga ggttttggc tatgaagggtc 1740
attacqatqq qatataatgtt qqqattacca atctactttc atqqaqtact qqqqqqatq 1800

ggatttgtt ttatggc ttggcgtgt ggagagtgt tggcgcgtat gtttttgtg 1860
aatcacgtca ttgagggtgt gagttatgga acgaaggatt tggtttgtgg tgcgagtc 1920
gtagatgaga agaagattgt caagccaaacg actgtattgg gagatacaccc aatggaaaag 1980
actcgcgagg aggcatgaa aagcaacagc aataacaaca agaagaaggg agagaagaac 2040
tcggtaccat ccgttccatt caacgactgg gcagcagtcc aatgccagac ctccgtaat 2100
tggctccag gctcatgggtt ctggaatcac tttctgggg gactctctca tcagattgag 2160
catcacttgt tccccagcat ttgtcataca aactactgtc atatccagga tgggtggag 2220
agtacgtgtg ctgagtgacgg agttccgtat cagagtgaga gtaattgtt tggcttat 2280
ggaaagatgat ttagtcattt gaagttttg ggtaaagcca agtgtgagta ggtgttaggt 2340
attgagaggt gtcgagttgt ctcatcttt aaaaataagc gctgaaagtg atttcgaaaa 2400
acaagggttgc tcaataccag tctctgtat tgattgtgc gtgcacacat ctccgtgagg 2460
agtttgacct cactcattct aacttggaat gtctctttg cgctggtgag ctggacgaa 2520
tacactccgn cagaagagac tgcatggta atgcagagga aagaggatat actgtatgag 2580
tccgaagaat cgatgacgacg cggtgagggtg gtgtacatca cttgtgagga ccaacgtgga 2640
accgcacatgtc tgaagagggtc catacctaaa catttgagcg gtctgggag caaactttag 2700
cagagattga atgctccattt cggtatttgtt tcattgtgc canattgtata aggaacagca 2760
accaacacac cqqqq 2775

<210> 13
<211> 564
<212> PRT
<213> Thalassiosira pseudonana

<400> 13

Met Cys Asn Gly Asn Leu Pro Ala Ser Thr Ala Gln Leu Lys Ser Thr
1 5 10 15

Ser Lys Pro Gln Gln Gln His Glu His Arg Thr Ile Ser Lys Ser Glu
20 25 30

Leu Ala Gln His Asn Thr Pro Lys Ser Ala Trp Cys Ala Val His Ser
35 40 45

Thr Pro Ala Thr Asp Pro Ser His Ser Asn Asn Lys Gln His Ala His
50 55 60

Leu Val Leu Asp Ile Thr Asp Phe Ala Ser Arg His Pro Gly Gly Asp
65 70 75 80

Leu Ile Leu Leu Ala Ser Gly Lys Asp Ala Ser Val Leu Phe Glu Thr
85 90 95

Tyr His Pro Arg Gly Val Pro Thr Ser Leu Ile Gln Lys Leu Gln Ile
100 105 110

Gly Val Met Glu Glu Ala Phe Arg Asp Ser Phe Tyr Ser Trp Thr
115 120 125

Asp Ser Asp Phe Tyr Thr Val Leu Lys Arg Arg Val Val Glu Arg Leu
130 135 140

Glu Glu Arg Gly Leu Ala Arg Arg Gly Ser Lys Glu Ile Trp Ile Lys
145 150 155 160

Ala Leu Phe Leu Leu Val Gly Phe Trp Tyr Cys Leu Tyr Lys Met Tyr
165 170 175

Thr Thr Ser Asp Ile Asp Gln Tyr Gly Ile Ala Ile Ala Tyr Ser Ile
180 185 190

Gly Met Gly Thr Phe Ala Ala Phe Ile Gly Thr Cys Ile Gln His Asp
195 200 205

Gly Asn His Gly Ala Phe Ala Gln Asn Lys Leu Leu Asn Lys Leu Ala
210 215 220

Gly Trp Thr Leu Asp Met Ile Gly Ala Ser Ala Phe Thr Trp Glu Leu
225 230 235 240

Gln His Met Leu Gly His His Pro Tyr Thr Asn Val Leu Asp Gly Val
245 250 255

Glu Glu Glu Arg Lys Glu Arg Gly Glu Asp Val Ala Leu Glu Glu Lys
260 265 270

Asp Gln Val Arg Arg Asp Asp Arg Glu Arg Glu Ser Leu Phe Val Gln
275 280 285

Glu Ser Asp Pro Asp Val Phe Ser Ser Phe Pro Leu Met Arg Met His
290 295 300

Pro Leu His Thr Thr Ser Trp Tyr His Lys Tyr Gln His Leu Tyr Ala
305 310 315 320

Pro Pro Leu Phe Ala Leu Met Thr Leu Ala Lys Val Phe Gln Gln Asp
325 330 335

Phe Glu Val Ala Thr Ser Gly Arg Leu Tyr His Ile Asp Ala Asn Val
340 345 350

Arg Tyr Gly Ser Val Trp Asn Val Met Arg Phe Trp Ala Met Lys Val
355 360 365

Ile Thr Met Gly Tyr Met Met Gly Leu Pro Ile Tyr Phe His Gly Val
370 375 380

Leu Arg Gly Val Gly Leu Phe Val Ile Gly His Leu Ala Cys Gly Glu
385 390 395 400

Leu Leu Ala Thr Met Phe Ile Val Asn His Val Ile Glu Gly Val Ser
405 410 415

Tyr Gly Thr Lys Asp Leu Val Gly Gly Ala Ser His Val Asp Glu Lys
420 425 430

Lys Ile Val Lys Pro Thr Thr Val Leu Gly Asp Thr Pro Met Glu Lys
435 440 445

Thr Arg Glu Glu Ala Leu Lys Ser Asn Ser Asn Asn Asn Lys Lys
450 455 460

Gly Glu Lys Asn Ser Val Pro Ser Val Pro Phe Asn Asp Trp Ala Ala
465 470 475 480

Val Gln Cys Gln Thr Ser Val Asn Trp Ser Pro Gly Ser Trp Phe Trp
485 490 495

Asn His Phe Ser Gly Gly Leu Ser His Gln Ile Glu His His Leu Phe
500 505 510

Pro Ser Ile Cys His Thr Asn Tyr Cys His Ile Gln Asp Val Val Glu
515 520 525

Ser Thr Cys Ala Glu Tyr Gly Val Pro Tyr Gln Ser Glu Ser Asn Leu
530 535 540

Phe Val Ala Tyr Gly Lys Met Ile Ser His Leu Lys Phe Leu Gly Lys
545 550 555 560

Ala Lys Cys Glu

<210> 14
<211> 1434
<212> DNA
<213> Thalassiosira pseudonana

<400> 14
atggacttgc tctccggcga tccttccgg acactcgcc ttgcagcact tgggttcatc 60
ggatttgctg cggcggtggca atgcttctac ccggccgagca tcgtcgccaa gcctcgatca 120
ttaagcaatg gttaaactcaa taccagaatc catggcaaat tgtacgacct ctcatcgtt 180
cagcatccag gaggccccgt ggctttctt cttgtcaag gtcgcgacgg aacagctcta 240
tttgagtcac accatccctt catacctcga aagaatctac ttcatcgatctt ctccaaatgtac 300
gagggtccgt cgactgaaga ctctgttcc ttcatcgcc cccttagacga actcaatgg 360
gaatctccgt acgatggaa ggacattgaa aatgtatgtt tcgtatctga cctacgagct 420
ctcgtaattt agcaactttc tcctctcgcc aaggaaaaggg gagtttcaact cggttgatcg 480
tcgaaggcaa cacctcagcg gtggatgggtt gttctactgc tccttcgtc gttttccctc 540

agcatcccat tatattttagg tggttcgtgg actttcggtg tcgtcactcc catcctcgct 600
tggctggcggttgcattactggcacatgactcactttgcatttgcagcaactgg 660
attttgcattatgcgtcccatatctcctccctctatcgagtccgtcaatgtggat 720
catcatcacttcattggaca tcacgcatac accaacatttccaaaagaga tccagatctt 780
gctcactcgctc cacaactcat gagagaacac aagagtatca aatggagacc atctcactta 840
aatcaaacac agctcccgcg gattcttc atctggtcga ttgcagtcgg tattgggtt 900
aacttactga acgacgtgag agcactaacc aagcttcat acaacaacgt tggcggtt 960
gagaagatgt catcgctcgcaacattactc catttccttg gacgtatgtt gcacatctt 1020
gtgactacac ttggccctt ttggcggtt ccgggtgtga aggccatcgtttgcact 1080
gtaccgaatgcataactgag ttgtgccttc atgctgaata cgcaaatcaa tcacccatc 1140
aacacgtgtc acatgcctc cgataacaac ttacaacgc atcaagtgtt aactgctcag 1200
aacttggcc gatcaagtgc ctggccttc atcttcgggaggctcaacttaccaaattt 1260
gaacatcatt tggccgac ggtgaaccat tgccatttgc cagcttgc cccgggtgtt 1320
gagcggtt gtaagaaaca cgggggtgaca tacaactctt tggaggata cagagaggcc 1380
atcattgcac acttgcaca taccaaagat atgtcgacga agcctactga ttga 1434

<210> 15
<211> 2906
<212> DNA
<213> Thalassiosira pseudonana

<220>
<221> misc_feature
<222> (2)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (14)..(14)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2645)..(2645)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2650)..(2650)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2655)..(2655)
<223> n is a, c, g, or t

<400> 15
anntctccca cccngccago tcttcaggt cgaccggaga tacacacttc ttcccaccaa 60
cttcgtccctc catacgatcg gaagaaaaga ggagattatc ttgacttctt gacggaggag 120
tggatgaaa agaacttgag tggtaaggg ctgatttcc tgagaaggag aagtcaatcg 180
gaacgaagtt catggagttt tggcaacc ctattgagac gttgcttgggtt ggagaaaggt 240

agcgagggtg agcatgcaaa cagaatggta taaatcacta agatgtcact cccaatgaca 300
agttagaata gcaatgacga gatgggtac agatgttaga gatggagaga ttaagcgaat 360
ggctggatga ttaggatatg caatgcaaaa ctgtatagat tcttgtaat agacttgta 420
gacaacgtcc gctgcagaa aaggacaata ctaattaata taaaaccgac tcggagagaa 480
catgacatgg caagttgtca ctatggatt cactacgtcg ctgacagga agtcacgtg 540
gcctcggcga agaagacaaa caaaaccgag ccctcacatt tcactctgta cagttcatag 600
tcaacaccac caatacgtatcccccaacg ccgatatctc ccgcacccgc aaccgcaccc 660
ccaccaaaaac aggtaccgtt gcctctgccc acaacaacga cccgcaccacc caatccgtcc 720
gaaccctcaa atctctcaag ggcaacgagg tcgtcatcaa cggcacaatt tatgacattt 780
ctgactttgt ccacatcgga ggagagggtt tcaagttctt tgggtggaaat gatgttacta 840
ttcagtataa tatgattcat ccgtatcata cggggaaaca tctggagaag atgaaggctg 900
ttggaaaggt ttagattgg cagtcggagt gagttgaat ggtgcacacg ttgacgttgt 960
tgggtgtca tttcggttccatcataatccactgatcataatccactacaatccgtt 1020
accatagcta caagttcgac acccccttg aacgagagat caaatcagaa gtgtcaaga 1080
tcgtacgtcg cgggcgtgag ttccggcaca caggctactt cctccgtgcc ttttctaca 1140
tcgctcttcttccatcataatccactgatcataatccactacaatccgtt 1200
atcactggta tcagagtggt gtattcatcg caattgtgtt tggatttca caggcatca 1260
ttgggtgaa tggccagcac gatgccaatc acggagctgc cagtaagcgt ccctgggtga 1320
atgacttggttt gggatttggaa acggatttga ttggatctaa caaatggaaat tggatggcac 1380
agcattggac tcatcagct tacactaacc atagtggaaat ggtcccgat agttcagct 1440
cggaacctat gttgcattc aatgactatc ccattggaca cccgaagaga aagtggggc 1500
ataggttcca gggagggtac ttcccttca tgctggact ttactggctc tcgactgtat 1560
tcaatccgca attcattgtatcataatccactgatcataatccactacaatccgtt 1620
agaatgattt cattgtcaag aggaggaagt acgcccgtgc attgaggatg atgtacattt 1680
acttgaacat tgcagcccc ttcatgaaca atggtttggat ctggctacc ttggaaatca 1740
tcatgttcatcataatccactgatcataatccactacaatccactacaatccgtt 1800
acttcatcaa ttccggatcgatcataatccactgatcataatccactacaatccgtt 1860
gctgggttcaatcataatccactgatcataatccactacaatccactacaatccgtt 1920
ttacggggagg actcaactt caggtggaaac atcatcttcccttccatcataatccactgatcataatccactacaatccgtt 1980
ggtatcccttcaatcataatccactgatcataatccactacaatccactacaatccgtt 2040
cttattatcc ttggatttggg cagaatttggatcataatccactgatcataatccactacaatccgtt 2100
gtatgtggcgc caactgggag ctcaagccgt tgcctggaaat ttagttgtac 2160
tgatgtcggtt gggatcgatcataatccactgatcataatccactacaatccactacaatccgtt 2220
cttgggtttt tggatcgatcataatccactgatcataatccactacaatccactacaatccgtt 2280
cctccacttg cactcaaaaac aaatcttagca taacatttcttccatcataatccactacaatccactacaatccgtt 2340

acgactcatt acgcaatgcc tctctataa ccccgaaaca actcgaccag cttcatactc 2400
taatcgcca tcctggcag ctgcaatcca gccctagcag cagctctt actcaactcc 2460
atcgactca acttcgtatc tgcccccga tcaatctat gcaaccgtgc cctcttacc 2520
aaatctgcct ttaacatcca gtaatcatag gcgattccac gtagtacgtt tgctcgctcg 2580
ggagacactg atgcccgtgc ttgttattgt gatatactgt gctggtgccg gcatcgatgc 2640
tccgntgtgn gttngactg tgcatggat gctgctgtga aacagtcggt gcagtgttagc 2700
ggaggtgctg ttctgaact gaggagatgc ccgcaaactg ataggggtg gtgcagcgct 2760
ataaaatttg cgagcgagtc cattgtccct gctctccca tatgtcgggc gagggcgaag 2820
cgcgaaggag aagccacaag gccaatacaa cagaaagttt aatgaagga cgtaattcct 2880
acacagtcca gtggcgaagt tacaac 2906

<210> 16
<211> 1434
<212> DNA
<213> Thalassiosira pseudonana

<400> 16
atggctcccc ccaacgcccga tatctcccgatccgcaacc gcatccccac caaaacaggt 60
accgttgccct ctggccgacaa caacgacccc gccaccataat ccgtccgaac cctcaaatct 120
ctcaaggcaca acgaggcgt catcaacggc acaatttatg acattgctga cttgtccat 180
cctggaggag aggttgtcaa gttctttgtt gggaaatgtt gttactattca gtataatatg 240
attcatccgt atcatacggg gaaacatctg gagaagatgtt aggcgttgg aaagggtgt 300
gattggcagt cgactacaa gttcgacacc cccttgaac gagagatcaa atcagaagt 360
ttcaagatcg tacgtcgccg gcgtgagttt ggcacaacag gctacttcctt ccgtgcctt 420
ttctacatcg ctcttctt caccatgcaat tacacttcg ccacatgcac caccatcacc 480
acctacgatc actggatca gagggtgtt ttcatcgcaat ttgtgttttgc tatttcacag 540
gcatttcattt ggttgaatgtt ccagcacat gccaatcacg gagctgccag taagcgtccc 600
tgggtgaatg acttgttggg atttggaaacg gatttgattt gatctaacaa atggaaattgg 660
atggcacacgc attggactca tcacgcitac actaaccata gtgagaagga tcccgatagc 720
ttcagctcgaa acctatgtt tgcatcaat gactatccca ttggacacccc gaagagaaag 780
tgggtggcata ggttcccgagg agggtacttc ctcttcattgc ttggacttta ctggctcccc 840
actgtattca atccgcaattt cattgtatccgtt cgtcaacgtt gggctcagta cgtcggaaattt 900
caaatggaga atgattcat tgcataaggagg aggaagtacg ccgttgcattt gaggatgtt 960
tacatttact tgaacattgtt cagccccatc atgaacaatgtt gtttgagctg gtctaccc 1020
ggaatcatca tggatgggg aatcagcgag agtctactc tcagttgtgtt cttctcggtt 1080
tctcacaact tcatcaattc ggatcgatcgtatc cctacggctg acttcaaaaaa gaccggagaa 1140
caagtgtgtt ggttcaagtc gcagggtggag acttcgttca cctatggggg ttttatttcc 1200
ggatgtcttca cgggaggactt caactttcag gtggaaacatc atctttcc cctgtatgagc 1260

agtgcgttgt atccttacat tgcacctacg gttcgtgagg tttgcaagaa gcacgggatg 1320
agctacgcctt attatccttg gattgggcag aatttggat caacattcaa atacatgcat 1380
cgcgctggta gtggagccaa ctgggagctc aagccgtgt ctggaaagtgc ctaa 1434

<210> 17
<211> 477
<212> PRT
<213> Thalassiosira pseudonana

<400> 17

Met Ala Pro Pro Asn Ala Asp Ile Ser Arg Ile Arg Asn Arg Ile Pro
1 5 10 15

Thr Lys Thr Gly Thr Val Ala Ser Ala Asp Asn Asn Asp Pro Ala Thr
20 25 30

Gln Ser Val Arg Thr Leu Lys Ser Leu Lys Gly Asn Glu Val Val Ile
35 40 45

Asn Gly Thr Ile Tyr Asp Ile Ala Asp Phe Val His Pro Gly Gly Glu
50 55 60

Val Val Lys Phe Phe Gly Gly Asn Asp Val Thr Ile Gln Tyr Asn Met
65 70 75 80

Ile His Pro Tyr His Thr Gly Lys His Leu Glu Lys Met Lys Ala Val
85 90 95

Gly Lys Val Val Asp Trp Gln Ser Asp Tyr Lys Phe Asp Thr Pro Phe
100 105 110

Glu Arg Glu Ile Lys Ser Glu Val Phe Lys Ile Val Arg Arg Gly Arg
115 120 125

Glu Phe Gly Thr Thr Gly Tyr Phe Leu Arg Ala Phe Phe Tyr Ile Ala
130 135 140

Leu Phe Phe Thr Met Gln Tyr Thr Phe Ala Thr Cys Thr Thr Phe Thr
145 150 155 160

Thr Tyr Asp His Trp Tyr Gln Ser Gly Val Phe Ile Ala Ile Val Phe
165 170 175

Gly Ile Ser Gln Ala Phe Ile Gly Leu Asn Val Gln His Asp Ala Asn
180 185 190

His Gly Ala Ala Ser Lys Arg Pro Trp Val Asn Asp Leu Leu Gly Phe
195 200 205

Gly Thr Asp Leu Ile Gly Ser Asn Lys Trp Asn Trp Met Ala Gln His
210 215 220

Trp Thr His His Ala Tyr Thr Asn His Ser Glu Lys Asp Pro Asp Ser
225 230 235 240

Phe Ser Ser Glu Pro Met Phe Ala Phe Asn Asp Tyr Pro Ile Gly His
245 250 255

Pro Lys Arg Lys Trp Trp His Arg Phe Gln Gly Gly Tyr Phe Leu Phe
260 265 270

Met Leu Gly Leu Tyr Trp Leu Pro Thr Val Phe Asn Pro Gln Phe Ile
275 280 285

Asp Leu Arg Gln Arg Gly Ala Gln Tyr Val Gly Ile Gln Met Glu Asn
290 295 300

Asp Phe Ile Val Lys Arg Arg Lys Tyr Ala Val Ala Leu Arg Met Met
305 310 315 320

Tyr Ile Tyr Leu Asn Ile Val Ser Pro Phe Met Asn Asn Gly Leu Ser
325 330 335

Trp Ser Thr Phe Gly Ile Ile Met Leu Met Gly Ile Ser Glu Ser Leu
340 345 350

Thr Leu Ser Val Leu Phe Ser Leu Ser His Asn Phe Ile Asn Ser Asp
355 360 365

Arg Asp Pro Thr Ala Asp Phe Lys Lys Thr Gly Glu Gln Val Cys Trp
370 375 380

Phe Lys Ser Gln Val Glu Thr Ser Ser Thr Tyr Gly Gly Phe Ile Ser
385 390 395 400

Gly Cys Leu Thr Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe
405 410 415

Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Thr Val Arg
420 425 430

Glu Val Cys Lys Lys His Gly Met Ser Tyr Ala Tyr Tyr Pro Trp Ile
435 440 445

Gly Gln Asn Leu Val Ser Thr Phe Lys Tyr Met His Arg Ala Gly Ser
450 455 460

Gly Ala Asn Trp Glu Leu Lys Pro Leu Ser Gly Ser Ala
465 470 475

<210> 18
<211> 1428
<212> DNA
<213> Thalassiosira pseudonana

<400> 18

<210> 19
<211> 475
<212> PRT
<213> Thalassiosira pseudonana

<400> 19

Met Ala Pro Pro Asn Ala Asp Ile Ser Arg Ile Arg Asn Arg Ile Pro
1 5 10 15

Thr Lys Thr Gly Thr Ser Ala Asp Asn Asn Asp Pro Ala Thr Gln Ser
20 25 30

Val Arg Thr Leu Lys Ser Leu Lys Gly Asn Glu Val Val Ile Asn Gly
35 40 45

Thr Ile Tyr Asp Ile Ala Asp Phe Val His Pro Gly Gly Glu Val Val
50 55 60

Lys Phe Phe Gly Gly Asn Asp Val Thr Ile Gln Tyr Asn Met Ile His
65 70 75 80

Pro Tyr His Thr Gly Lys His Leu Glu Lys Met Lys Ala Val Gly Lys
85 90 95

Val Val Asp Trp Gln Ser Asp Tyr Lys Phe Asp Thr Pro Phe Glu Arg
100 105 110

Glu Ile Lys Ser Glu Val Phe Lys Ile Val Arg Arg Gly Arg Glu Phe
115 120 125

Gly Thr Thr Gly Tyr Phe Leu Arg Ala Phe Phe Tyr Ile Ala Leu Phe
130 135 140

Phe Thr Met Gln Tyr Thr Phe Ala Thr Cys Thr Thr Phe Thr Thr Tyr
145 150 155 160

Asp His Trp Tyr Gln Ser Gly Val Phe Ile Ala Ile Val Phe Gly Ile
165 170 175

Ser Gln Ala Phe Ile Gly Leu Asn Val Gln His Asp Ala Asn His Gly
180 185 190

Ala Ala Ser Lys Arg Pro Trp Val Asn Asp Leu Leu Gly Phe Gly Thr
195 200 205

Asp Leu Ile Gly Ser Asn Lys Trp Asn Trp Met Ala Gln His Trp Thr
210 215 220

His His Ala Tyr Thr Asn His Ser Glu Lys Asp Pro Asp Ser Phe Ser
225 230 235 240

Ser Glu Pro Met Phe Ala Phe Asn Asp Tyr Pro Ile Gly His Pro Lys
245 250 255

Arg Lys Trp Trp His Arg Phe Gln Gly Gly Tyr Phe Leu Phe Met Leu
260 265 270

Gly Leu Tyr Trp Leu Ser Thr Val Phe Asn Pro Gln Phe Ile Asp Leu
275 280 285

Arg Gln Arg Gly Ala Gln Tyr Val Gly Ile Gln Met Glu Asn Asp Phe
290 295 300

Ile Val Lys Arg Arg Lys Tyr Ala Val Ala Leu Arg Met Met Tyr Ile
305 310 315 320

Tyr Leu Asn Ile Val Ser Pro Phe Met Asn Asn Gly Leu Ser Trp Ser
325 330 335

Thr Phe Gly Ile Ile Met Leu Met Gly Ile Ser Glu Ser Leu Thr Leu
340 345 350

Ser Val Leu Phe Ser Leu Ser His Asn Leu Ile Asn Ser Asp Arg Asp
355 360 365

Pro Thr Ala Asp Phe Lys Lys Thr Gly Glu Gln Val Cys Trp Phe Lys
370 375 380

Ser Gln Val Glu Thr Ser Ser Thr Tyr Gly Gly Phe Ile Ser Gly Cys
385 390 395 400

Leu Thr Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe Pro Arg
405 410 415

Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Thr Val Arg Glu Val
420 425 430

Cys Lys Lys His Gly Val Asn Tyr Ala Tyr Tyr Pro Trp Ile Gly Gln
435 440 445

Asn Leu Val Ser Thr Phe Lys Tyr Met His Arg Ala Gly Ser Gly Ala
450 455 460

Asn Trp Glu Leu Lys Pro Leu Ser Gly Ser Ala
465 470 475

<210> 20
<211> 1485
<212> DNA
<213> Thalassiosira pseudonana

<400> 20
atggctccac cctccatcaa agacacactc gacgagccct tcgtctcgcc cgcattccacc 60
aagtgcggca ccaccaaacc cctccctccc cgccgcaaac ccctcaaacg atactccccc 120
tcccaaatct cccaacacaa cactccccacc gatgcatggc tcattacaa atccaaatgc 180
cttgacattt ccaaattggat atgcacccat ccagggtggag agcagacgct gttggaggtt 240
gccggatgg atgctaccga tgaattgagg gcattcaig atgattgggt ttggaggag 300
aagtgcctc attttgtat tggggagggtg gattggacta ctacggcg ggcagagaat 360
actgtcacga aggatggaca ggttcggag cttatcaagg atttcagaga gttgggtgaa 420
cacttcgaca ggttggggta ctttcacgtc agtccatggt attacgtccg taagggtggct 480
accgtcttcg ccatcttgg atgtgcactc ggactcctct tcaataccga ttccatccca 540
gcacacatgc tcgcggcggt actcctcggt atattctggc aacaattgc attcgtcgga 600
catgactgtg gtcacatgtc ggcgcggact catgcccgtg atcatatcga tgtacctaag 660
ctgggagcac tggtgaccctt cttcaatggg atttcggtag cgtggtgaa ggctacgcac 720
aatgttcatc atgtgtgcc aaatagtgtt gattgtgacc cggacattgc tcatttgcgg 780
gtgttgcgt tgcatgagca catgttacg tcgttgtta acaagtatca tgggagggtg 840

atggagttt attggctggc gcgtaatgtc ttgtgccat ttcaacactt ttggactat 900
ccataatgg cggggcgag gttcaatctg tacattcaat cagcattgtt ttggcgctcg 960
aagaacgatg ggcattcagg aagaacaaca ttggattga tggcgatcat cggcttc 1020
tcttggtagt cgggtcggt gtcattcatc ccgagctggc cggagcgatcgatcattcg 1080
ttcgctagcc atgctgtac tggttactg aatgtcaaa tcacactgtc gcacttct 1140
cggccaatct ttgataccaa caaagaggga cccaggttt gagggtactt ttactctcg 1200
aacgtccttg cttcggttga cgtcgctgt cctacatact tggactgggtt ccacggaggt 1260
ctccaattcc aaacactcca tcattgtac cctagacttg gacgtcagca ctggagaaag 1320
accgaaccc tcattgtac gttgtcaag aagcattttt taccatacac gagcaagagc 1380
ttcgtagagt gcaatatgg agttttaat acattgaagg atgcccgcgcg ttctgccaag 1440
aagtggtcac cgttaattta tgagtcaatg tgtgctcagg gatag 1485

<210> 21
<211> 494
<212> PRT
<213> Thalassiosira pseudonana

<400> 21

Met Ala Pro Pro Ser Ile Lys Asp Thr Leu Asp Glu Pro Phe Val Ser
1 5 10 15

Pro Ala Ser Thr Lys Ser Pro Thr Thr Lys Pro Leu Leu Pro Arg Arg
20 25 30

Lys Pro Leu Lys Arg Tyr Ser Pro Ser Gln Ile Ser Gln His Asn Thr
35 40 45

Pro Thr Asp Ala Trp Leu Ile Tyr Lys Ser Gln Val Leu Asp Ile Ser
50 55 60

Lys Trp Ile Ser His His Pro Gly Gly Glu Gln Thr Leu Leu Arg Phe
65 70 75 80

Ala Gly Met Asp Ala Thr Asp Glu Leu Arg Ala Phe His Asp Asp Trp
85 90 95

Val Leu Glu Glu Lys Leu Pro His Phe Val Ile Gly Glu Val Asp Trp
100 105 110

Thr Thr Thr Gly Gly Ala Glu Asn Thr Val Thr Lys Asp Gly Gln Val
115 120 125

Ser Glu Leu Ile Lys Asp Phe Arg Glu Leu Gly Glu His Phe Asp Arg
130 135 140

Leu Gly Tyr Phe His Val Ser Pro Trp Tyr Tyr Val Arg Lys Val Ala
145 150 155 160

Thr Val Phe Ala Ile Phe Gly Cys Ala Leu Gly Leu Leu Phe Asn Thr
165 170 175

Asp Ser Ile Pro Ala His Met Leu Ala Ala Val Leu Leu Gly Ile Phe
180 185 190

Trp Gln Gln Phe Ala Phe Val Gly His Asp Cys Gly His Met Ser Ala
195 200 205

Arg Thr His Ala Arg Asp His Ile Asp Val Pro Lys Leu Gly Ala Leu
210 215 220

Val Thr Phe Phe Asn Gly Ile Ser Val Ala Trp Trp Lys Ala Thr His
225 230 235 240

Asn Val His His Ala Val Pro Asn Ser Val Asp Cys Asp Pro Asp Ile
245 250 255

Ala His Leu Pro Val Phe Ala Leu His Glu His Met Phe Thr Ser Leu
260 265 270

Phe Asn Lys Tyr His Gly Arg Val Met Glu Phe Asp Trp Leu Ala Arg
275 280 285

Asn Val Phe Val Pro Phe Gln His Phe Trp Tyr Tyr Pro Ile Met Ala
290 295 300

Val Ala Arg Phe Asn Leu Tyr Ile Gln Ser Ala Leu Phe Leu Ala Ser
305 310 315 320

Lys Asn Asp Gly His Ala Gly Arg Thr Thr Leu Asp Leu Met Ala Phe
325 330 335

Ile Gly Phe Phe Ser Trp Leu Ala Val Leu Val Ser Cys Ile Pro Ser
340 345 350

Trp Pro Glu Arg Ile Ala Phe Val Phe Val Ser His Ala Val Ala Gly
355 360 365

Leu Leu Asn Val Gln Ile Thr Leu Ser His Phe Ser Arg Pro Ile Phe
370 375 380

Asp Thr Asn Lys Glu Gly Pro Arg Phe Gly Gly Asp Phe Tyr Ser Arg
385 390 395 400

Asn Val Leu Ala Ser Leu Asp Val Ala Cys Pro Thr Tyr Leu Asp Trp
405 410 415

Phe His Gly Gly Leu Gln Phe Gln Thr Leu His His Cys Tyr Pro Arg
420 425 430

Leu Gly Arg Gln His Leu Arg Lys Thr Glu Pro Leu Ile Ala Ser Leu
435 440 445

Cys Lys Lys His Ser Leu Pro Tyr Thr Ser Lys Ser Phe Val Glu Cys
450 455 460

Asn Met Glu Val Phe Asn Thr Leu Lys Asp Ala Ala Arg Ser Ala Lys
465 470 475 480

Lys Trp Ser Pro Leu Ile Tyr Glu Ser Met Cys Ala Gln Gly
485 490

<210> 22
<211> 16
<212> DNA
<213> Thalassiosira pseudonana

<400> 22
ggtaacgaat tggtag 16

<210> 23
<211> 16
<212> DNA
<213> Thalassiosira pseudonana

<400> 23
gtcggcatag tttatg 16

<210> 24
<211> 21
<212> DNA
<213> Thalassiosira pseudonana

<400> 24
gtgagagcac taaccaagct t 21

<210> 25
<211> 19
<212> DNA
<213> Thalassiosira pseudonana

<400> 25
caatcgttag gcttcgtcg 19

<210> 26
<211> 33
<212> DNA
<213> Thalassiosira pseudonana

<400> 26
gcgggatcca ccatggctgg aaaaggagga gac 33

<210> 27
<211> 26
<212> DNA
<213> Thalassiosira pseudonana

<400> 27
gcgaattctt acatggcagg gaaatc 26

<210> 28
<211> 33

<212> DNA

<213> Thalassiosira pseudonana

<400> 28

gcgggatcca ccatggctga ctttctctcc ggc

33

<210> 29

<211> 26

<212> DNA

<213> Thalassiosira pseudonana

<400> 29

gcgaattctc aatcagtagg ctgcgt

26

<210> 30

<211> 444

<212> PRT

<213> Danio rerio

<400> 30

Met Gly Gly Gly Gln Gln Thr Asp Arg Ile Thr Asp Thr Asn Gly
1 5 10 15

Arg Phe Ser Ser Tyr Thr Trp Glu Glu Val Gln Lys His Thr Lys His
20 25 30

Gly Asp Gln Trp Val Val Val Glu Arg Lys Val Tyr Asn Val Ser Gln
35 40 45

Trp Val Lys Arg His Pro Gly Gly Leu Arg Ile Leu Gly His Tyr Ala
50 55 60

Gly Glu Asp Ala Thr Glu Ala Phe Thr Ala Phe His Pro Asn Leu Gln
65 70 75 80

Leu Val Arg Lys Tyr Leu Lys Pro Leu Leu Ile Gly Glu Leu Glu Ala
85 90 95

Ser Glu Pro Ser Gln Asp Arg Gln Lys Asn Ala Ala Leu Val Glu Asp
100 105 110

Phe Arg Ala Leu Arg Glu Arg Leu Glu Ala Glu Gly Cys Phe Lys Thr
115 120 125

Gln Pro Leu Phe Phe Ala Leu His Leu Gly His Ile Leu Leu Leu Glu
130 135 140

Ala Ile Ala Phe Met Met Val Trp Tyr Phe Gly Thr Gly Trp Ile Asn
145 150 155 160

Thr Leu Ile Val Ala Val Ile Leu Ala Thr Ala Gln Ser Gln Ala Gly
165 170 175

Trp Leu Gln His Asp Phe Gly His Leu Ser Val Phe Lys Thr Ser Gly
180 185 190

Met Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
195 200 205

Ser Ala Gly Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
210 215 220

Asn Ile Phe Lys Lys Asp Pro Asp Val Asn Met Leu Asn Ala Phe Val
225 230 235 240

Val Gly Asn Val Gln Pro Val Glu Tyr Gly Val Lys Lys Ile Lys His
245 250 255

Leu Pro Tyr Asn His Gln His Lys Tyr Phe Phe Ile Gly Pro Pro
260 265 270

Leu Leu Ile Pro Val Tyr Phe Gln Phe Ile Phe His Asn Met Ile
275 280 285

Ser His Gly Met Trp Val Asp Leu Leu Trp Cys Ile Ser Tyr Tyr Val
290 295 300

Arg Tyr Phe Leu Cys Tyr Thr Gln Phe Tyr Gly Val Phe Trp Ala Ile
305 310 315 320

Ile Leu Phe Asn Phe Val Arg Phe Met Glu Ser His Trp Phe Val Trp
325 330 335

Val Thr Gln Met Ser His Ile Pro Met Asn Ile Asp Tyr Glu Lys Asn
340 345 350

Gln Asp Trp Leu Ser Met Gln Leu Val Ala Thr Cys Asn Ile Glu Gln
355 360 365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
370 375 380

His His Leu Phe Pro Thr Val Pro Arg His Asn Tyr Trp Arg Ala Ala
385 390 395 400

Pro Arg Val Arg Ala Leu Cys Glu Lys Tyr Gly Val Lys Tyr Gln Glu
405 410 415

Lys Thr Leu Tyr Gly Ala Phe Ala Asp Ile Ile Arg Ser Leu Glu Lys
420 425 430

Ser Gly Glu Leu Trp Leu Asp Ala Tyr Leu Asn Lys
435 440

<210> 31

<211> 444

<212> PRT

<213> Homo sapiens

<400> 31

Met Gly Lys Gly Gly Asn Gln Gly Glu Gly Ala Ala Glu Arg Glu Val
1 5 10 15

Ser Val Pro Thr Phe Ser Trp Glu Glu Ile Gln Lys His Asn Leu Arg
20 25 30

Thr Asp Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Thr Lys
35 40 45

Trp Ser Ile Gln His Pro Gly Gly Gln Arg Val Ile Gly His Tyr Ala
50 55 60

Gly Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Pro Asp Leu Glu
65 70 75 80

Phe Val Gly Lys Phe Leu Lys Pro Leu Leu Ile Gly Glu Leu Ala Pro
85 90 95

Glu Glu Pro Ser Gln Asp His Gly Lys Asn Ser Lys Ile Thr Glu Asp
100 105 110

Phe Arg Ala Leu Arg Lys Thr Ala Glu Asp Met Asn Leu Phe Lys Thr
115 120 125

Asn His Val Phe Phe Leu Leu Leu Ala His Ile Ile Ala Leu Glu
130 135 140

Ser Ile Ala Trp Phe Thr Val Phe Tyr Phe Gly Asn Gly Trp Ile Pro
145 150 155 160

Thr Leu Ile Thr Ala Phe Val Leu Ala Thr Ser Gln Ala Gln Ala Gly
165 170 175

Trp Leu Gln His Asp Tyr Gly His Leu Ser Val Tyr Arg Lys Pro Lys
180 185 190

Trp Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
195 200 205

Ser Ala Asn Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
210 215 220

Asn Ile Phe His Lys Asp Pro Asp Val Asn Met Leu His Val Phe Val
225 230 235 240

Leu Gly Glu Trp Gln Pro Ile Glu Tyr Gly Lys Lys Lys Leu Lys Tyr
245 250 255

Leu Pro Tyr Asn His Gln His Glu Tyr Phe Phe Leu Ile Gly Pro Pro
260 265 270

Leu Leu Ile Pro Met Tyr Phe Gln Tyr Gln Ile Ile Met Thr Met Ile
275 280 285

Val His Lys Asn Trp Val Asp Leu Ala Trp Ala Val Ser Tyr Tyr Ile
290 295 300

Arg Phe Phe Ile Thr Tyr Ile Pro Phe Tyr Gly Ile Leu Gly Ala Leu
305 310 315 320

Leu Phe Leu Asn Phe Ile Arg Phe Leu Glu Ser His Trp Phe Val Trp
325 330 335

Val Thr Gln Met Asn His Ile Val Met Glu Ile Asp Gln Glu Ala Tyr
340 345 350

Arg Asp Trp Phe Ser Ser Gln Leu Thr Ala Thr Cys Asn Val Glu Gln
355 360 365

Ser Phe Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
370 375 380

His His Leu Phe Pro Thr Met Pro Arg His Asn Leu His Lys Ile Ala
385 390 395 400

Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Glu
405 410 415

Lys Pro Leu Leu Arg Ala Leu Leu Asp Ile Ile Arg Ser Leu Lys Lys
420 425 430

Ser Gly Lys Leu Trp Leu Asp Ala Tyr Leu His Lys
435 440

<210> 32
<211> 444
<212> PRT
<213> Homo sapiens

<400> 32

Met Ala Pro Asp Pro Val Ala Ala Glu Thr Ala Ala Gln Gly Pro Thr
1 5 10 15

Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln Arg Ser Gly Cys Glu
20 25 30

Glu Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Ser Glu Phe
35 40 45

Thr Arg Arg His Pro Gly Gly Ser Arg Val Ile Ser His Tyr Ala Gly
50 55 60

Gln Asp Ala Thr Asp Pro Phe Val Ala Phe His Ile Asn Lys Gly Leu
65 70 75 80

Val Lys Lys Tyr Met Asn Ser Leu Leu Ile Gly Glu Leu Ser Pro Glu
85 90 95

Gln Pro Ser Phe Glu Pro Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe
100 105 110

Arg Glu Leu Arg Ala Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn
115 120 125

His Val Phe Phe Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly
130 135 140

Ala Ala Trp Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe
145 150 155 160

Leu Leu Cys Ala Val Leu Leu Ser Ala Val Gln Ala Gln Ala Gly Trp
165 170 175

Leu Gln His Asp Phe Gly His Leu Ser Val Phe Ser Thr Ser Lys Trp
180 185 190

Asn His Leu Leu His His Phe Val Ile Gly His Leu Lys Gly Ala Pro
195 200 205

Ala Ser Trp Trp Asn His Met His Phe Gln His His Ala Lys Pro Asn
210 215 220

Cys Phe Arg Lys Asp Pro Asp Ile Asn Met His Pro Phe Phe Phe Ala
225 230 235 240

Leu Gly Lys Ile Leu Ser Val Glu Leu Gly Lys Gln Lys Lys Tyr
245 250 255

Met Pro Tyr Asn His Gln His Lys Tyr Phe Phe Leu Ile Gly Pro Pro
260 265 270

Ala Leu Leu Pro Leu Tyr Phe Gln Trp Tyr Ile Phe Tyr Phe Val Ile
275 280 285

Gln Arg Lys Lys Trp Val Asp Leu Ala Trp Met Ile Thr Phe Tyr Val
290 295 300

Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu
305 310 315 320

Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp
325 330 335

Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn
340 345 350

Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys
Page 41

355

360

365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
370 375 380

His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Lys Val Ala
385 390 395 400

Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser
405 410 415

Lys Pro Leu Leu Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu
420 425 430

Ser Gly Gln Leu Trp Leu Asp Ala Tyr Leu His Gln
435 440

<210> 33

<211> 439

<212> PRT

<213> Thraustochytrium sp.

<400> 33

Met Gly Lys Gly Ser Glu Gly Arg Ser Ala Ala Arg Glu Met Thr Ala
1 5 10 15

Glu Ala Asn Gly Asp Lys Arg Lys Thr Ile Leu Ile Glu Gly Val Leu
20 25 30

Tyr Asp Ala Thr Asn Phe Lys His Pro Gly Gly Ser Ile Ile Asn Phe
35 40 45

Leu Thr Glu Gly Glu Ala Gly Val Asp Ala Thr Gln Ala Tyr Arg Glu
50 55 60

Phe His Gln Arg Ser Gly Lys Ala Asp Lys Tyr Leu Lys Ser Leu Pro
65 70 75 80

Lys Leu Asp Ala Ser Lys Val Glu Ser Arg Phe Ser Ala Lys Glu Gln
85 90 95

Ala Arg Arg Asp Ala Met Thr Arg Asp Tyr Ala Ala Phe Arg Glu Glu
100 105 110

Leu Val Ala Glu Gly Tyr Phe Asp Pro Ser Ile Pro His Met Ile Tyr
115 120 125

Arg Val Val Glu Ile Val Ala Leu Phe Ala Leu Ser Phe Trp Leu Met
130 135 140

Ser Lys Ala Ser Pro Thr Ser Leu Val Leu Gly Val Val Met Asn Gly
145 150 155 160

Ile Ala Gln Gly Arg Cys Gly Trp Val Met His Glu Met Gly His Gly
165 170 175

Ser Phe Thr Gly Val Ile Trp Leu Asp Asp Arg Met Cys Glu Phe Phe
180 185 190

Tyr Gly Val Gly Cys Gly Met Ser Gly His Tyr Trp Lys Asn Gln His
195 200 205

Ser Lys His His Ala Ala Pro Asn Arg Leu Glu His Asp Val Asp Leu
210 215 220

Asn Thr Leu Pro Leu Val Ala Phe Asn Glu Arg Val Val Arg Lys Val
225 230 235 240

Lys Pro Gly Ser Leu Leu Ala Leu Trp Leu Arg Val Gln Ala Tyr Leu
245 250 255

Phe Ala Pro Val Ser Cys Leu Leu Ile Gly Leu Gly Trp Thr Leu Tyr
260 265 270

Leu His Pro Arg Tyr Met Leu Arg Thr Lys Arg His Met Glu Phe Val
275 280 285

Trp Ile Phe Ala Arg Tyr Ile Gly Trp Phe Ser Leu Met Gly Ala Leu
290 295 300

Gly Tyr Ser Pro Gly Thr Ser Val Gly Met Tyr Leu Cys Ser Phe Gly
305 310 315 320

Leu Gly Cys Ile Tyr Ile Phe Leu Gln Phe Ala Val Ser His Thr His
325 330 335

Leu Pro Val Thr Asn Pro Glu Asp Gln Leu His Trp Leu Glu Tyr Ala
340 345 350

Ala Asp His Thr Val Asn Ile Ser Thr Lys Ser Trp Leu Val Thr Trp
355 360 365

Trp Met Ser Asn Leu Asn Phe Gln Ile Glu His His Leu Phe Pro Thr
370 375 380

Ala Pro Gln Phe Arg Phe Lys Glu Ile Ser Pro Arg Val Glu Ala Leu
385 390 395 400

Phe Lys Arg His Asn Leu Pro Tyr Tyr Asp Leu Pro Tyr Thr Ser Ala
405 410 415

Val Ser Thr Thr Phe Ala Asn Leu Tyr Ser Val Gly His Ser Val Gly
420 425 430

Ala Asp Thr Lys Lys Gln Asp
435

<210> 34
<211> 443
<212> PRT
<213> *Caenorhabditis elegans*

<400> 34

Met Val Val Asp Lys Asn Ala Ser Gly Leu Arg Met Lys Val Asp Gly
1 5 10 15

Lys Trp Leu Tyr Leu Ser Glu Glu Leu Val Lys Lys His Pro Gly Gly
20 25 30

Ala Val Ile Glu Gln Tyr Arg Asn Ser Asp Ala Thr His Ile Phe His
35 40 45

Ala Phe His Glu Gly Ser Ser Gln Ala Tyr Lys Gln Leu Asp Leu Leu
50 55 60

Lys Lys His Gly Glu His Asp Glu Phe Leu Glu Lys Gln Leu Glu Lys
65 70 75 80

Arg Leu Asp Lys Val Asp Ile Asn Val Ser Ala Tyr Asp Val Ser Val
85 90 95

Ala Gln Glu Lys Lys Met Val Glu Ser Phe Glu Lys Leu Arg Gln Lys
100 105 110

Leu His Asp Asp Gly Leu Met Lys Ala Asn Glu Thr Tyr Phe Leu Phe
115 120 125

Lys Ala Ile Ser Thr Leu Ser Ile Met Ala Phe Ala Phe Tyr Leu Gln
130 135 140

Tyr Leu Gly Trp Tyr Ile Thr Ser Ala Cys Leu Leu Ala Leu Ala Trp
145 150 155 160

Gln Gln Phe Gly Trp Leu Thr His Glu Phe Cys His Gln Gln Pro Thr
165 170 175

Lys Asn Arg Pro Leu Asn Asp Thr Ile Ser Leu Phe Phe Gly Asn Phe
180 185 190

Leu Gln Gly Phe Ser Arg Asp Trp Trp Lys Asp Lys His Asn Thr His
195 200 205

His Ala Ala Thr Asn Val Ile Asp His Asp Gly Asp Ile Asp Leu Ala
210 215 220

Pro Leu Phe Ala Phe Ile Pro Gly Asp Leu Cys Lys Tyr Lys Ala Ser
225 230 235 240

Phe Glu Lys Ala Ile Leu Lys Ile Val Pro Tyr Gln His Leu Tyr Phe
Page 44

245

250

255

Thr Ala Met Leu Pro Met Leu Arg Phe Ser Trp Thr Gly Gln Ser Val
260 265 270

Gln Trp Val Phe Lys Glu Asn Gln Met Glu Tyr Lys Val Tyr Gln Arg
275 280 285

Asn Ala Phe Trp Glu Gln Ala Thr Ile Val Gly His Trp Ala Trp Val
290 295 300

Phe Tyr Gln Leu Phe Leu Leu Pro Thr Trp Pro Leu Arg Val Ala Tyr
305 310 315 320

Phe Ile Ile Ser Gln Met Gly Gly Leu Leu Ile Ala His Val Val
325 330 335

Thr Phe Asn His Asn Ser Val Asp Lys Tyr Pro Ala Asn Ser Arg Ile
340 345 350

Leu Asn Asn Phe Ala Ala Leu Gln Ile Leu Thr Thr Arg Asn Met Thr
355 360 365

Pro Ser Pro Phe Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln Ile
370 375 380

Glu His His Leu Phe Pro Thr Met Pro Arg Cys Asn Leu Asn Ala Cys
385 390 395 400

Val Lys Tyr Val Lys Glu Trp Cys Lys Glu Asn Asn Leu Pro Tyr Leu
405 410 415

Val Asp Asp Tyr Phe Asp Gly Tyr Ala Met Asn Leu Gln Gln Leu Lys
420 425 430

Asn Met Ala Glu His Ile Gln Ala Lys Ala Ala
435 440

<210> 35

<211> 443

<212> PRT

<213> *Caenorhabditis elegans*

<400> 35

Met Val Val Asp Lys Asn Ala Ser Gly Leu Arg Met Lys Val Asp Gly
1 5 10 15

Lys Trp Leu Tyr Leu Ser Glu Glu Leu Val Lys Lys His Pro Gly Gly
20 25 30

Ala Val Ile Glu Gln Tyr Arg Asn Ser Asp Ala Thr His Ile Phe His
35 40 45

Ala Phe His Glu Gly Ser Ser Gln Ala Tyr Lys Gln Leu Asp Leu Leu
50 55 60

Lys Lys His Gly Glu His Asp Glu Phe Leu Glu Lys Gln Leu Glu Lys
65 70 75 80

Arg Leu Asp Lys Val Asp Ile Asn Val Ser Ala Tyr Asp Val Ser Val
85 90 95

Ala Gln Glu Lys Lys Met Val Glu Ser Phe Glu Lys Leu Arg Gln Lys
100 105 110

Leu His Asp Asp Gly Leu Met Lys Ala Asn Glu Thr Tyr Phe Leu Phe
115 120 125

Lys Ala Ile Ser Thr Leu Ser Ile Met Ala Phe Ala Phe Tyr Leu Gln
130 135 140

Tyr Leu Gly Trp Tyr Ile Thr Ser Ala Cys Leu Leu Ala Leu Ala Trp
145 150 155 160

Gln Gln Phe Gly Trp Leu Thr His Glu Phe Cys His Gln Gln Pro Thr
165 170 175

Lys Asn Arg Pro Leu Asn Asp Thr Ile Ser Leu Phe Phe Gly Asn Phe
180 185 190

Leu Gln Gly Phe Ser Arg Asp Trp Trp Lys Asp Lys His Asn Thr His
195 200 205

His Ala Ala Thr Asn Val Ile Asp His Asp Gly Asp Ile Asp Leu Ala
210 215 220

Pro Leu Phe Ala Phe Ile Pro Gly Asp Leu Cys Lys Tyr Lys Ala Ser
225 230 235 240

Phe Glu Lys Ala Ile Leu Lys Ile Val Pro Tyr Gln His Leu Tyr Phe
245 250 255

Thr Ala Met Leu Pro Met Leu Arg Phe Ser Trp Thr Gly Gln Ser Val
260 265 270

Gln Trp Val Phe Lys Glu Asn Gln Met Glu Tyr Lys Val Tyr Gln Arg
275 280 285

Asn Ala Phe Trp Glu Gln Ala Thr Ile Val Gly His Trp Ala Trp Val
290 295 300

Phe Tyr Gln Leu Phe Leu Leu Pro Thr Trp Pro Leu Arg Val Ala Tyr
305 310 315 320

Phe Ile Ile Ser Gln Met Gly Gly Leu Leu Ile Ala His Val Val
325 330 335

Thr Phe Asn His Asn Ser Val Asp Lys Tyr Pro Ala Asn Ser Arg Ile
340 345 350

Leu Asn Asn Phe Ala Ala Leu Gln Ile Leu Thr Thr Arg Asn Met Thr
355 360 365

Pro Ser Pro Phe Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln Ile
370 375 380

Glu His His Leu Phe Pro Thr Met Pro Arg Cys Asn Leu Asn Ala Cys
385 390 395 400

Val Lys Tyr Val Lys Glu Trp Cys Lys Glu Asn Asn Leu Pro Tyr Leu
405 410 415

Val Asp Asp Tyr Phe Asp Gly Tyr Ala Met Asn Leu Gln Gln Leu Lys
420 425 430

Asn Met Ala Glu His Ile Gln Ala Lys Ala Ala
435 440

<210> 36

<211> 419

<212> PRT

<213> Euglena gracilis

<400> 36

Met Lys Ser Lys Arg Gln Ala Leu Ser Pro Leu Gln Leu Met Glu Gln
1 5 10 15

Thr Tyr Asp Val Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu
20 25 30

Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met His Phe
35 40 45

Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Pro Ser
50 55 60

Phe Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe
65 70 75 80

Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser
85 90 95

Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly Leu Gly Val
100 105 110

Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe Ile Gly Ala
115 120 125

Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp
Page 47

130

135

140

Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Leu Val
145 150 155 160

Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Cys Trp
165 170 175

Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His
180 185 190

Asp Pro Asp Ile Asp Asn Leu Pro Pro Leu Ala Trp Ser Glu Asp Asp
195 200 205

Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln
210 215 220

Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe
225 230 235 240

Gln Cys Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe
245 250 255

Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His Trp
260 265 270

Thr Leu Lys Ala Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr
275 280 285

Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly Ile
290 295 300

Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly Asp
305 310 315 320

Pro Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu Thr
325 330 335

Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly Leu
340 345 350

Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn
355 360 365

Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His Asn
370 375 380

Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu Leu
385 390 395 400

Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln Pro Ala Gly
405 410 415

Lys Ala Leu

<210> 37

<211> 477

<212> PRT

<213> Phaeodactylum tricornutum

<400> 37

Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala
1 5 10 15

Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
20 25 30

Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
35 40 45

Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
50 55 60

Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
65 70 75 80

Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
85 90 95

Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
100 105 110

Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
115 120 125

Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
130 135 140

Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
145 150 155 160

Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
165 170 175

His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe
180 185 190

Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys
195 200 205

His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val
210 215 220

Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp
225 230 235 240

Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys
245 250 255

Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr
260 265 270

Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe
275 280 285

Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu
290 295 300

Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile
305 310 315 320

Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg
325 330 335

Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser
340 345 350

Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met
355 360 365

Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val
370 375 380

Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe
385 390 395 400

Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu
405 410 415

Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val
420 425 430

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu
435 440 445

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
450 455 460

Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
465 470 475

<210> 38

<211> 404

<212> PRT

<213> Borago officinalis

<400> 38

Ser Phe Pro Leu Lys Ser Leu Ala Gly Gln Glu Val Thr Asp Ala Phe

1 5 10 15

Val Ala Phe His Pro Ala Ser Thr Trp Lys Asn Leu Asp Lys Phe Phe
20 25 30

Thr Gly Tyr Tyr Leu Lys Asp Tyr Ser Val Ser Glu Val Ser Lys Asp
35 40 45

Tyr Arg Lys Leu Val Phe Glu Phe Ser Lys Met Gly Leu Tyr Asp Lys
50 55 60

Lys Gly His Ile Met Phe Ala Thr Leu Cys Phe Ile Ala Met Leu Phe
65 70 75 80

Ala Met Ser Val Tyr Gly Val Leu Phe Cys Glu Gly Val Leu Val His
85 90 95

Leu Phe Ser Gly Cys Leu Met Gly Phe Leu Trp Ile Gln Ser Gly Trp
100 105 110

Ile Gly His Asp Ala Gly His Tyr Met Val Val Ser Asp Ser Arg Leu
115 120 125

Asn Lys Phe Met Gly Ile Phe Ala Ala Asn Cys Leu Ser Gly Ile Ser
130 135 140

Ile Gly Trp Trp Lys Trp Asn His Asn Ala His His Ile Ala Cys Asn
145 150 155 160

Ser Leu Glu Tyr Asp Pro Asp Leu Gln Tyr Ile Pro Phe Leu Val Val
165 170 175

Ser Ser Lys Phe Phe Gly Ser Leu Thr Ser His Phe Tyr Glu Lys Arg
180 185 190

Leu Thr Phe Asp Ser Leu Ser Arg Phe Phe Val Ser Tyr Gln His Trp
195 200 205

Thr Phe Tyr Pro Ile Met Cys Ala Ala Arg Leu Asn Met Tyr Val Gln
210 215 220

Ser Leu Ile Met Leu Leu Thr Lys Arg Asn Val Ser Tyr Arg Ala His
225 230 235 240

Glu Leu Leu Gly Cys Leu Val Phe Ser Ile Trp Tyr Pro Leu Leu Val
245 250 255

Ser Cys Leu Pro Asn Trp Gly Glu Arg Ile Met Phe Val Ile Ala Ser
260 265 270

Leu Ser Val Thr Gly Met Gln Gln Val Gln Phe Ser Leu Asn His Phe
275 280 285

Ser Ser Ser Val Tyr Val Gly Lys Pro Lys Gly Asn Asn Trp Phe Glu
290 295 300

Lys Gln Thr Asp Gly Thr Leu Asp Ile Ser Cys Pro Pro Trp Met Asp
305 310 315 320

Trp Phe His Gly Gly Leu Gln Phe Gln Ile Glu His His Leu Phe Pro
325 330 335

Lys Met Pro Arg Cys Asn Leu Arg Lys Ile Ser Pro Tyr Val Ile Glu
340 345 350

Leu Cys Lys His Asn Leu Pro Tyr Asn Tyr Ala Ser Phe Ser Lys
355 360 365

Ala Asn Glu Met Thr Leu Arg Thr Leu Arg Asn Thr Ala Leu Gln Ala
370 375 380

Arg Asp Ile Thr Lys Pro Leu Pro Lys Asn Leu Val Trp Glu Ala Leu
385 390 395 400

His Thr His Gly